Thesis/
Reports
Shepard,
B. B.

FINAL REPORT FOR RESEARCH AGREEMENT
#INT-93845-RJVA
"Trout Population Structure"
MONTANA STATE UNIVERSITY
FS Contact: Dr. Bruce Rieman
CoOp Contact: Dr. Robert White

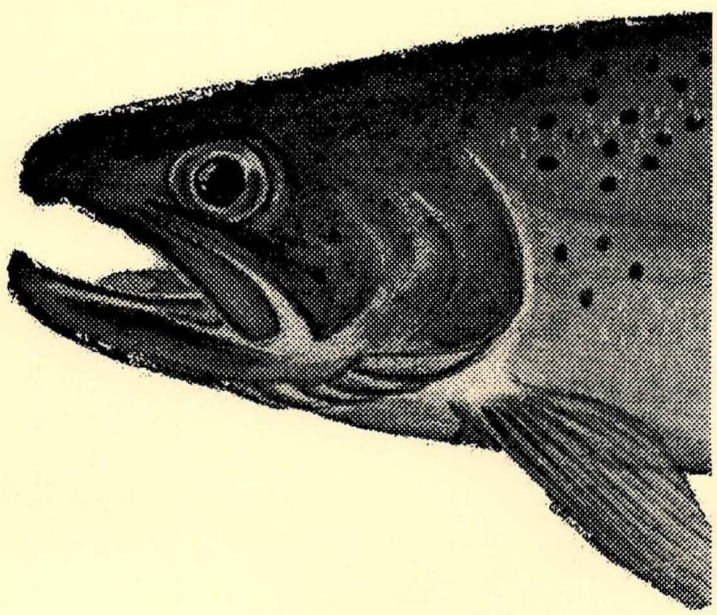
Movement and population structure of westslope cutthroat trout *Oncorhynchus clarki lewisi* inhabiting headwaters streams of Montana

Final Report to:
USDA, Forest Service, Rocky Mountain Research Station
316 East Myrtle Street
Boise, Idaho 83702

Contract: INT-93845-RJVA

1998

Montana Cooperative Fishery
Research Unit
Biology Department
Montana State University
Bozeman, Montana 59717



Cover illustration adapted from a color painting by Glenn West

FINAL REPORT FOR RESEARCH AGREEMENT
#INT-93845-RJVA
"Trout Population Structure"
MONTANA STATE UNIVERSITY
FS Contact: Dr. Bruce Rieman
CoOp Contact: Dr. Robert White

Movement and population structure of westslope cutthroat trout *Oncorhynchus clarki lewisi* inhabiting headwaters streams of Montana

Bradley B. Shepard
Montana Department of Fish, Wildlife and Parks
Department of Biology
Montana State University, Bozeman, Montana 59717, USA

Jim Robison-Cox
Department of Mathematical Sciences
Montana State University, Bozeman, Montana 59717, USA

and

Susan C. Ireland and Robert G. White
Montana Cooperative Fishery Research Unit[a]
Department of Biology, Montana State University
Bozeman, Montana 59717, USA

Final Report to:
USDA, Forest Service, Rocky Mountain Research Station
316 East Myrtle Street
Boise, Idaho 83702

Contract: **INT-93845-RJVA**

August 1998

[a] The Unit is jointly sponsored by Montana State University, Montana Department of Fish, Wildlife, and Parks, and the U.S. National Biological Service.

Movement and population structure of westslope cutthroat trout *Oncorhynchus clarki lewisi* inhabiting headwaters streams of Montana

Bradley B. Shepard
Montana Department of Fish, Wildlife and Parks
Department of Biology
Montana State University, Bozeman, Montana 59717, USA

Jim Robison-Cox
Department of Mathematical Sciences
Montana State University, Bozeman, Montana 59717, USA

and

Susan C. Ireland and Robert G. White
Montana Cooperative Fishery Research Unit[a]
Department of Biology, Montana State University
Bozeman, Montana 59717, USA

Final Report to:
USDA, Forest Service, Rocky Mountain Research Station
316 East Myrtle Street
Boise, Idaho 83702

Contract: **INT-93845-RJVA**

August 1998

[a] The Unit is jointly sponsored by Montana State University, Montana Department of Fish, Wildlife, and Parks, and the U.S. National Biological Service.

Executive Summary

Retention of visible implant (VI) tags in westslope cutthroat trout Oncorhynchus clarki lewisi inhabiting 11 isolated headwater tributary drainages in Montana was evaluated during 1993 and 1994. In 1993 2,071 VI tags were implanted in westslope cutthroat trout (100 to 324 mm; FL) and adipose fins were removed as a secondary mark to evaluate tag retention. Of 348 westslope cutthroat trout recaptured during the year they were tagged, 201 (58%) had retained their tags. Of 616 westslope cutthroat trout recaptured the year after tagging, 355 (58%) had retained their tags. Logistic regression analyses indicated that fish length was the most significant variable that positively influenced tag retention. Other significant variables were wetted width and channel gradient of the stream in which fish were tagged, and quality of tag insertion, which was rated at time of tagging. Fish condition did not significantly improve deviance performance of logistic regression models that included fork length and tag insertion quality. Neither slopes nor intercepts of $\log_{10}$length-$\log_{10}$weight regressions were significantly different ($P > 0.10$) for fish which retained versus lost their tags. Fish condition was not significantly different ($P > 0.95$l; ANCOVA) between previously tagged and untagged westslope cutthroat trout after accounting for differences between drainages and years. We found no significant differences between slopes ($P > 0.50$) or intercepts ($P > 0.05$) of $\log_{10}$length-$\log_{10}$weight regressions between previously tagged and untagged fish. However, for 11 drainages where comparisons could be made, we found significant differences ($P < 0.05$) between $\log_{10}$length-$\log_{10}$weight regression slopes for previously tagged versus untagged fish in one drainage and for regression intercepts in an additional three drainages. Ninety-four percent of all tags were readable at recapture. A logistic regression model predicted that tag retention would be 75% or higher in westslope cutthroat trout $\geq$ 155 mm (FL) if tag insertion quality was good. In spite of relatively poor tag retention (< 75%) in smaller (< 155 mm) westslope cutthroat trout, VI tags were a valuable tool to assess movements of those fish, which retained their tags.

Movement of resident westslope cutthroat trout was also assessed using VI tags from 1993 to 1995. Little movement over 500 m was observed; however, the proportion of fish moving further than 0.5 km appeared to be correlated to level of physical isolation of that population. Length-frequency histograms of captured westslope cutthroat trout could be used for identifying missing year-classes, especially the year immediately following the emergence of the missing year-class. Our ability to interpret age classes from length-frequency histograms was inconsistent among streams. For some streams we could interpret age classes 1, 2, and sometimes, 3 from histograms, but for some others we could not. We could not discern any age classes beyond age 3 from histograms. We recommend a consistent time of year be used as a standard for collecting length-frequency data for comparative purposes, and suggest late summer for this purpose.

Table of Contents

Executive Summary ii
Table of Contents iii
List of Tables iv
List of Figures v
Introduction 1
Chapter 1: Factors influencing retention of visible implant tags in westslope cutthroat trout *Oncorhynchus clarki lewisi* inhabiting headwater streams of Montana 2
Abstract 3
Introduction 4
Study Area 4
Methods 4
Results 11
Discussion 12
Chapter 2: Movement of resident westslope cutthroat trout inhabiting isolated headwater streams of Montana 16
Introduction 16
Study Area 17
Methods 17
Results 18
Discussion 23
Chapter 3: Length frequencies of westslope cutthroat trout resident of headwater streams in Montana 24
Introduction 24
Results and Discussion 24
Acknowledgements 31
References 32

List of Tables

Table 1. Physical features of streams (numbers in parentheses are map locations shown in Figure 1) where densities of westslope cutthroat trout were estimated from 1992 to 1995. Mean temperatures, conductivities, and pH's are averages for several point estimates taken during field sampling. . "Model sample size" refers to number of fish tagged in each reach (see "Methods" for further explanation). 6

Table 2. Results of weighted logistic regression analyses to assess the probability of retention of visible implant tags in westslope cutthroat trout showing the improvement in model deviance the last entered explanatory variable provided. Explanatory variables are fork length at tagging (FL), tag insertion quality rating (R), wetted stream width (W), stream channel gradient as a percentage (G), and drainage (LOC). Explanatory variable coefficients and their associated standard errors (in parentheses) are shown for each model in variable order starting with the intercept, where the last explanatory variable added significantly to the model. Model run only for tagging data where tag insertions were rated. 13

Table 3. Ranks assigned to levels of isolation for westslope cutthroat trout populations where movement was evaluated from 1993 to 1995. 18

Table 4. Length (km) of habitat occupied by westslope cutthroat trout and length (%) of occupied habitats sampled by season and year (Su93 = Summer 1993, Fa93 = Fall 1993, Sp93 = Spring 1993, Su94 = Summer 1995, and Su95 = Summer 1995) by stream. 19

Introduction

Spatial diversity and dispersal among populations of animals is thought to play an important role in regional or metapopulation persistence. Metapopulation processes have not been clearly demonstrated in interior salmonids and the scale of population structuring is unknown. The general objective of this project was to improve our understanding of the movements of trout at a local level. The two specific objectives were:

1. Describe the movement and population structure of westslope cutthroat trout within an individual stream.
2. Determine whether habitat condition influences dispersal among stream reaches and habitat units.

In this final report we present a paper we published in the North American Journal of Fisheries Management in 1996 (Volume 16: 913-920) that documents what factors influenced the retention of visible implant tags in westslope cutthroat trout as Chapter 1. A companion paper authored by Jim Robison-Cox of Montana State University (Robison-Cox 1998) appears in Appendix A. In Chapter 2 we present summaries of tag recovery data showing movement patterns observed in different stream populations. In Chapter 3 we present length-frequency summaries by drainage by year. The data collected during this study is available in electronic format. The descriptions of these data are found in Appendix B. Two companion reports were prepared. One report details the "Population dynamics and demographics of westslope cutthroat trout Oncorhynchus clarki lewisi inhabiting isolated headwater tributaries of Montana" (Shepard et al. 1998a). The other describes the "Influence of abiotic and biotic factors on abundance of stream-resident westslope cutthroat trout Oncorhynchus clarki lewisi in Montana streams" (Shepard et al. 1998b).

Chapter 1:
Factors influencing retention of visible implant tags in westslope cutthroat trout *Oncorhynchus clarki lewisi* inhabiting headwater streams of Montana

Bradley B. Shepard
Montana Department of Fish, Wildlife and Parks
Department of Biology
Montana State University, Bozeman, Montana 59717, USA

Jim Robison-Cox
Department of Mathematical Sciences
Montana State University, Bozeman, Montana 59717, USA

and

Susan C. Ireland and Robert G. White
Montana Cooperative Fishery Research Unit[a]
Department of Biology, Montana State University
Bozeman, Montana 59717, USA

[a] The Unit is jointly sponsored by Montana State University, Montana Department of Fish, Wildlife, and Parks, and the U.S. National Biological Service.

Abstract

Retention of visible implant (VI) tags in westslope cutthroat trout Oncorhynchus clarki lewisi inhabiting 20 reaches of 11 isolated headwater tributary drainages in Montana was evaluated during 1993 and 1994. In 1993 2,071 VI tags were implanted in westslope cutthroat trout (100 to 324 mm; FL) and adipose fins were removed as a secondary mark to evaluate tag retention. Of 348 westslope cutthroat trout recaptured during the year they were tagged, 201 (58%) had retained their tags. Of 616 westslope cutthroat trout recaptured the year after tagging, 355 (58%) had retained their tags. Logistic regression analyses indicated that fish length was the most significant variable that positively influenced tag retention. Other significant variables were wetted width and channel gradient of the stream in which fish were tagged, and quality of tag insertion, which was rated at time of tagging. Fish condition did not significantly improve deviance performance of logistic regression models that included fork length and tag insertion quality. Neither slopes nor intercepts of $\log_{10}$length-$\log_{10}$weight regressions were significantly different ($P > 0.10$) for fish which retained versus lost their tags. Fish condition was not significantly different ($P > 0.951$; ANCOVA) between previously tagged and untagged westslope cutthroat trout after accounting for differences between drainages and years. We found no significant differences between slopes ($P > 0.50$) or intercepts ($P > 0.05$) of $\log_{10}$length-$\log_{10}$weight regressions between previously tagged and untagged fish. However, for 11 drainages where comparisons could be made, we found significant differences ($P < 0.05$) between $\log_{10}$length-$\log_{10}$weight regression slopes for previously tagged versus untagged fish in one drainage and for regression intercepts in an additional three drainages. Ninety-four percent of all tags were readable at recapture. A logistic regression model predicted that tag retention would be 75% or higher in westslope cutthroat trout $\geq$ 155 mm (FL) if tag insertion quality was good. In spite of relatively poor tag retention (< 75%) in smaller (< 155 mm) westslope cutthroat trout, VI tags were a valuable tool to assess movements of those fish, which retained their tags.

Introduction

Fish researchers and managers mark fish to obtain information on abundance, movements and migration, age and growth, mortality, behavior, exploitation rates, and stocking success (McFarlane et al. 1990). Evaluating retention of marks is important in any mark-recapture study (e.g., Nielsen 1992). Haw et al. (1990) developed an alphanumeric coded visible implant (VI) tag that can be inserted just beneath clear tissue, usually in a post-orbital location, and read upon subsequent recapture. Due to their small size, VI tags can be used to uniquely identify small fish without sacrificing the fish to recover the tag. Initial testing of VI tags by Haw et al. (1990) documented that only one tag was lost from 42 tagged rainbow trout (Oncorhynchus mykiss; 149-172 mm TL; mean 158 mm) held 22 weeks in a hatchery environment.

Visible implant tag retention has been related to fish size (length and/or weight), or age, for Atlantic salmon Salmo salar (Kincaid and Calkins 1992), brook trout Salvelinus fontinalis (Bryan and Ney 1994), and brown trout Salmo trutta (Niva 1995). McMahon et al. (1996) suggested that environmental conditions might also affect tag retention due to observed differences in tag retention between rainbow trout raised in a hatchery versus wild environment. We needed to document the retention of VI tags as part of a study to estimate mortality and movement of small (100-320 mm, FL) westslope cutthroat trout Oncorhynchus clarki lewisi inhabiting headwater streams of Montana. We wanted to determine what factors influenced tag loss and to test if VI tags adversely affected the condition of tagged fish.

Study Area

Visible implant tag retention was assessed in westslope cutthroat trout from 20 reaches within 11 tributary drainages of the upper Missouri River and two tributary drainages of the upper Clark Fork River, Montana (Figure 1). These streams were characterized by low-flow wetted widths of 1.5 to 5.8 m, water conductivities of 72 to 661 umhos, elevations of 1,320 to 2,570 m, and gradients of 3 to 11% (Table 1). Westslope cutthroat trout in sampled populations exhibited little evidence of introgression with either rainbow trout or Yellowstone cutthroat trout Oncorhynchus clarki bouveri using external morphometric characteristics and/or electrophoretic testing (Salmon and Trout Genetics Laboratory, University of Montana, Missoula, Montana, unpublished data).

Methods

Fish were captured using a Smith-Root BP-15 backpack electrofisher set at 40 mHz pulse with a pulse width of 1 msec. Voltages were set from 400 to 700 volts. These settings were used to maximize capture efficiency while simultaneously minimizing risk of injuring fish (W. Fredenberg, U.S. Fish and Wildlife Service, personal communication). All captured fish were anesthetized with MS-222, measured to the nearest millimeter (FL), and weighed to the nearest gram. Condition factors ($weight*10^5/length^3$) were calculated for individual fish (Anderson and Gutreuter 1983).

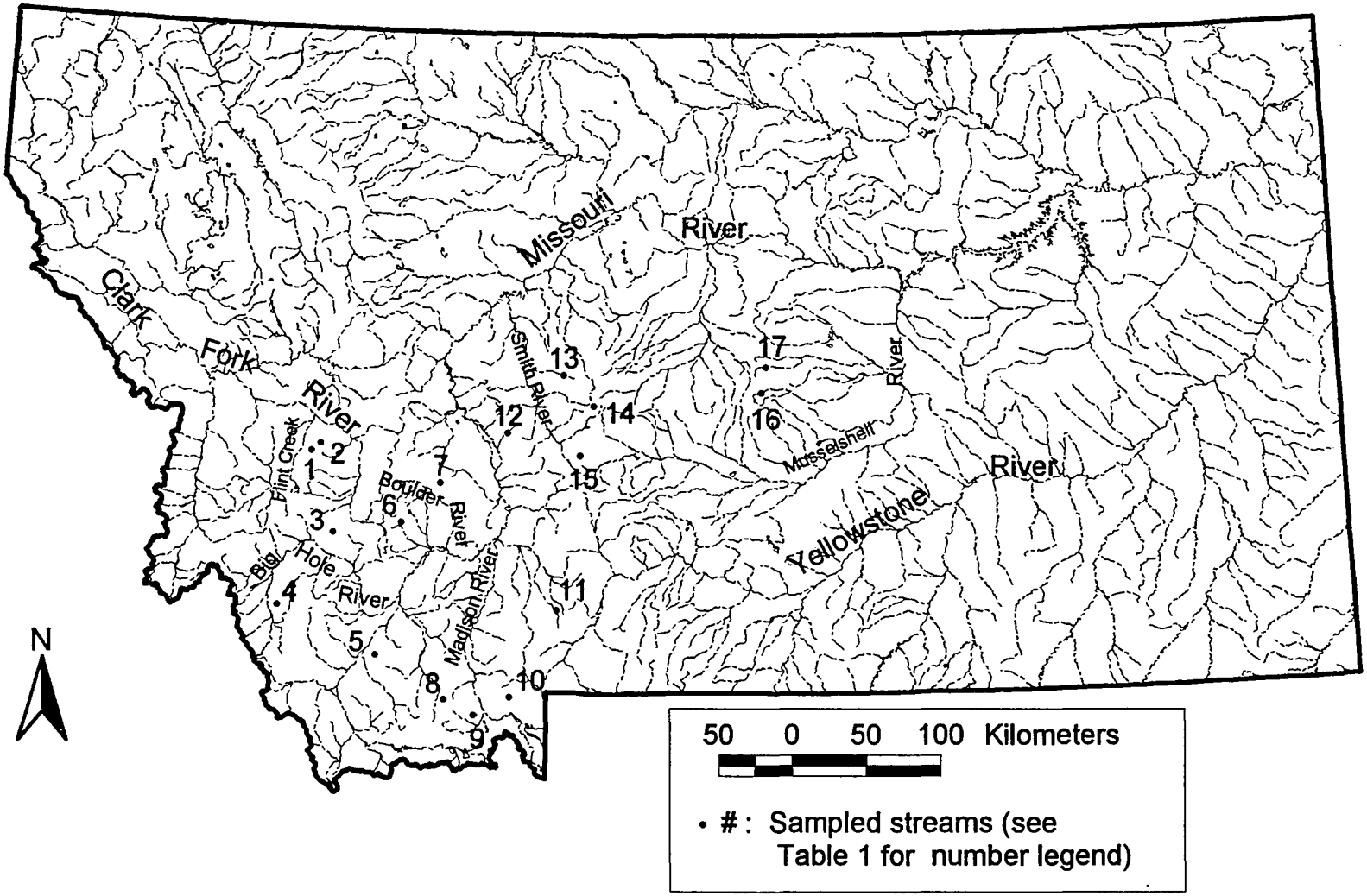


Figure 1. Map of Montana showing sample locations (dots) where tag retention, movement, and size structure of westslope cutthroat trout was assessed. Numbers next to locations correspond to numbers in parentheses by stream names in Table 1. Only sites 1, 2, 3, 6, 8, 10, 12, 14, 15, 16, and 17 were used for assessing tag retention in 1993 and 1994.

Table 1. Physical features of streams (numbers in parentheses are map locations shown in Figure 1) where densities of westslope cutthroat trout were estimated from 1992 to 1995. Mean temperatures, conductivities, and pH's are averages for several point estimates taken during field sampling. . "Model sample size" refers to number of fish tagged in each reach (see "Methods" for further explanation).

Stream (map number)	Rosgen[a] Type	Mean temperature (F)	Mean Conductivity (μmhos)	pH	Mean wetted width (m)	Length with cutthroat (km)	Lower elevation (m)	Upper elevation (m)	Mean channel gradient (%)	Number of tags out	Model sample size
Cabin (10)	C3	57.2	329	8.8	6.6	2.9	2170	2304	4.6	-[c]	-
M Fork Cabin	B3	53.0	203	8.7	3.3	7.0	2300	2600	4.2	-	-
Unnamed Trib	B3	46.6	268	8.7	1.6	1.0	2540	2560	2.4	-	-
Collar Gulch (17)	B2	46.2	174	8.3	2.0	2.7	1450	1550	3.6	232	134
Cottonwood (Smith)(15)	B1	46.0	130	8.9	2.5	0.6	1830	1850	4.0	89	88
East Fork Cottonwood	A3	48.8	144	8.8	1.5	3.0	1850	2190	11.4	25	0
West Fork Cottonwood	A3	47.9	126	8.1	2.5	2.8	1850	2110	9.1	333	124
Douglas (1)	A2	41.5	245	8.9	3.0	1.1	1570	1630	5.5	81	26
N. Fork Douglas (1)	B2	47.1	252	8.5	1.5	3.1	1650	1790	4.7	152	39
Geyser (8)	A3	48.2	380	8.6	1.6	1.6	2460	2570	6.6	305	79
Half Moon (16)	A2	49.9	344	8.7	3.0	7.3	1710	2010	4.2	11	0
Halfway (6)	B3	49.2	72	8.8	1.7	7.8	1830	2290	5.9	193	73
Jerry (3)	A3	43.7	156	8.7	5.8	2.8	2050	2220	3.0	95	28

Table 1. (Continued).

Stream (map number)	Rosgen[a] Type	Mean temperature (F)	Mean Conductivity (μmhos)	pH	Mean wetted width (m)	Length with cutthroat (km)	Lower elevation (m)	Upper elevation (m)	Mean channel gradient (%)	Number of tags out	Model sample size
Lick (11)	A3	51.1	262	8.4	2.3	2.0	1950	2020	6.1	-	-
Mcvey (4)	B3	42.9	64	9.3	1.5	2.3	1860	1940	3.4	-	-
Muskrat (7)	A2	46.0	NA[b]	NA[b]	3.6	2.2	1570	1700	5.6	-	-
N. Fork Deadman (14)	A3	43.3	235	8.6	1.7	2.5	1960	2130	6.8	5	0
N. Fork Gold (2)	B3	44.3	166	8.5	1.9	4.2	1880	2010	3.2	102	16
Soap Creek(10)	A3	46.7	72	8.5	2.5	3.9	1910	2170	6.6	166	58
Stone Creek (5)											
Left Fork	A3	50.0	296	8.0	2.1	3.5	1920	2150	6.6	-	-
Middle Fork	B4	51.0	336	8.0	1.4	1.2	1920	1970	4.6	-	-
Tenderfoot (13)	A3	44.0	116	8.6		8.1	1730	3000	4.5	-	-
White's Gulch (12)	B3	47.5	661	8.2	1.5	4.6	1320	1470	3.3	38	29

[a] Rosgen channel types based on Rosgen (1994).
[b] Data "Not Available".
[c] Tag retention study not done in these streams..

All westslope cutthroat trout from 120 mm to the largest fish captured (324 mm) and a few between 100 and 120 mm were tagged with individually-coded alpha-numeric VI tags as described by Haw et al. (1990) and Kincaid and Calkins (1992)(Figure 2). A minimum fish length of 150 mm (TL) is recommended by the manufacturer to obtain reasonable tag retention (P. Bergman, Northwest Marine Technology, personal communication); however, we were interested in obtaining information on movements and retention of tags for smaller fish. We implanted 2,071 rectangular, plastic-laminate tags (2.5 mm x 1.0 mm x 0.1 mm thick) with a white code on a black background in 1993. Most fish were tagged behind the left eye where tags were injected into the clear post-orbital adipose tissue with a Northwest Marine Technology tagging syringe. A few fish were tagged behind the right eye when tag insertion was poor behind the left eye. Each person who tagged fish had extensive experience applying VI tags (at least 1,000 tags) prior to this study. Adipose fins were excised from all tagged fish to assess tag retention. Prior to excision, no westslope cutthroat trout were missing their adipose fin in any of our sampled streams.

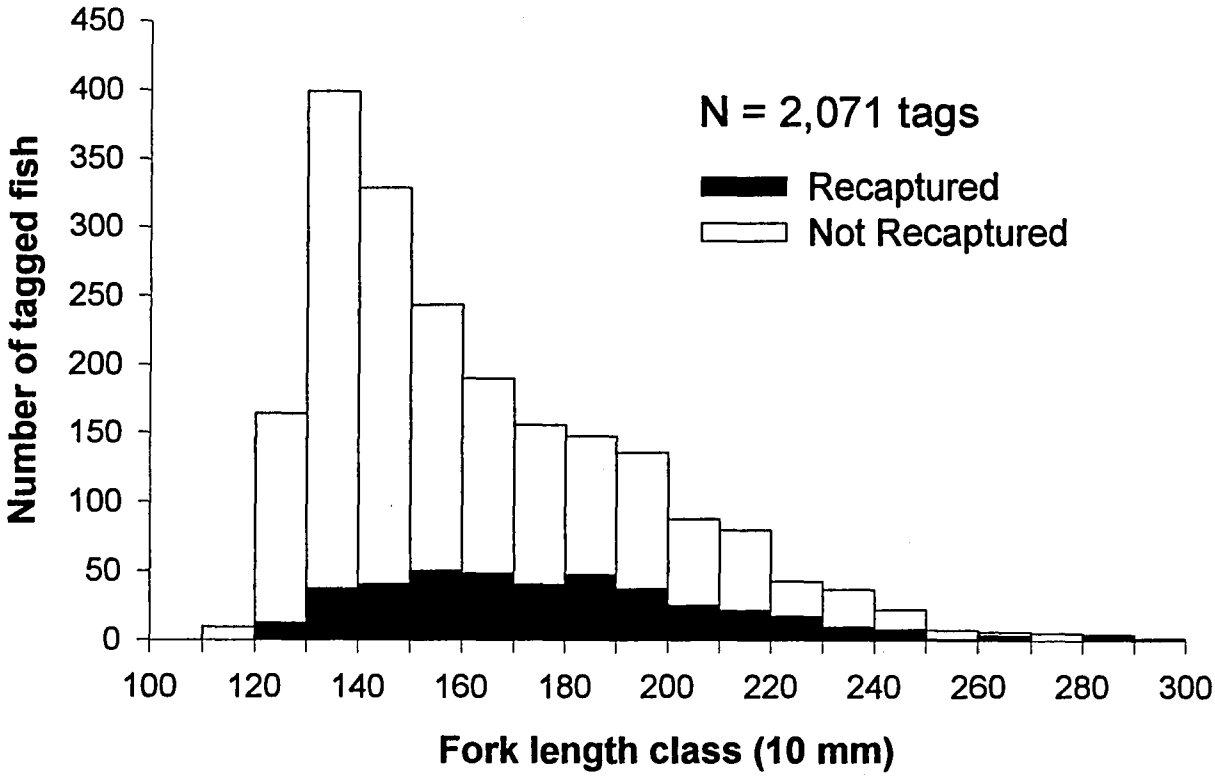


Figure 2. Size distribution of westslope cutthroat trout tagged with visible implant tags in 1993 and of those recaptured with tags during 1993 and 1994.

We rated tagging quality using three classes: 1) good - tag properly inserted without tearing any adipose tissue and easily read; 2) fair - some adipose tissue torn or tag inserted too deep for easy reading; or 3) poor - large rip in adipose tissue or tag inserted so deeply that it was extremely difficult to read or could not be read. Of the 2,071 tags implanted, 1,421 (69%) were classed as "good" insertions, 494 (24%) as "fair", and 34 (1%) as "poor", while 122 (6%) were not classed. Following tagging, fish were allowed to fully recover from the effects of the anesthetic in a perforated bucket prior to being released into calm water areas in the stream.

Recoveries of tagged fish were made from 1 to 406 d after tagging. We followed the protocol of Geoghegan et al. (1990) for quality control and assurance of tag data with the following modifications. When a fish with a clipped adipose fin was recaptured, an attempt was made to find and read the VI tag by inspecting post-orbital tissue behind each eye. If two separate people could not see a VI tag it was recorded as lost. Those fish that had lost their original VI tag were re-tagged, if possible. We were unable to re-tag eight adipose clipped fish in 1993 and 35 in 1994.

Upon recapture, the tag's code was recorded when possible. If a tag was difficult to read, a second person read the tag in an attempt to reduce observer error. Tag readability was classified as unreadable, difficult to read (there was a question on one or more of the letters or numbers), or readable. All retained tags were considered as tag recaptures, regardless of readability. Tag retention at recapture was estimated within the year tagged and the year following tagging by dividing the number of recaptured westslope cutthroat trout with tags by the number recaptured with adipose fin clips.

McMahon et al. (1996) and P. Bergman (Northwest Marine Technology, personal communication) indicated that different stocks inhabiting different environments might retain VI tags differentially. To assess potential effects of different stream environments on tag retention using easily measured variables, wetted stream widths and reach gradients were estimated. Wetted stream widths were measured and averaged to the nearest 0.1 m for each reach. Reach gradients were estimated to the nearest percent using 1:24,000 U.S. Geological Survey contour maps. Reaches were sub-divided based on distribution of westslope cutthroat trout and changes in channel gradient (Table 1). Tag retention was coded as 0 for missing or 1 for present so that logistic regression (Hosmer and Lemeshow 1989) could be used to model the probability of tag retention. We used weighted logistic regression where a weighting variable was assigned to each tagged fish based on our estimated probability that the fish was subsequently recaptured. A weighting variable of 1 was assigned to each fish subsequently recaptured with a tag. Because fin clips did not uniquely identify each fish, length at tagging was unknown for fish that had lost their tags.

To estimate length at tagging for recaptured fish that had lost their tags; we first estimated daily growth by tributary drainage based on tag-recaptured fish and an assumed growth season of May 15 to September 30. We then estimated the 95% prediction interval for length at tagging by back-calculating from length at recapture using daily growth estimates for each tributary

drainage and time-at-large for each recaptured fish which had lost its tag. Regression weightings were assigned to all tagged fish that were not subsequently recaptured with a retained tag by assessing the probability that they were recaptured with a lost tag. We did that by examining 95% prediction intervals for estimated length at tagging of fish recaptured with lost tags during the next sampling event. We attempted to match the recaptured fish known to have lost tags with the fish that were tagged and not recaptured. For this matching we assumed fish did not move more than 200 m between tagging and first recapture, which was true for more than 95% of the tag-recaptured fish.

For example, if a fish was recaptured without a tag we could estimate that its length at tagging was between 165 to 185 mm (95% prediction interval) using the growth regression. All fish (n) of unknown fate (ie., not recaptured with a tag) tagged within 200 m of the recapture location with a recorded length at tagging within the predicted 165-185 mm length interval were assigned weightings according to the formula: Weighting = 1/n. If three fish fit the criteria, then each was given a weighting factor of 0.33. We followed this procedure for each adipose marked fish recaptured without a tag and weightings accumulated to a maximum of one for each fish tagged.

For weighted logistic regression we had a sample size of 770 fish (see Table 1 for sample size by reach). We were unable to weigh 58 fish, consequently for models that tested effects of fish condition, and our sample size was 712, which included 359 fish recaptured with tags and 353 fish of unknown fate. The effective degrees of freedom were 501. We did not rate tag insertion quality for 59 fish, consequently models which tested effects of tag insertion quality had a sample size of 711 with 500 effective degrees of freedom. Explanatory variables explored were fork length at tagging (FL), condition factor (C), tag rating (R), reach gradient (G), wetted width (W), and drainage as a class variable (LOC). Variables were added to the model one at a time and tested to determine if they added significantly to the model by comparing the two models under a Chi-square distribution with the appropriate degrees of freedom (Hosmer and Lemeshow 1989).

We tested for differences in length-weight relationships between those fish which lost tags and those fish which retained tags; and between fish which had, to those that had not, been previously tagged by comparing $\log_{10}$length-$\log_{10}$weight regressions (Anderson and Gutreuter 1983). If regression slopes were found to not be significantly different, we tested for differences in elevations (intercepts) according to methods described by Zar (1974). We also tested for influences of tagging on fish condition by comparing condition factors between previously tagged and untagged cutthroat trout $\geq$ 110 mm using ANCOVA and by comparing $\log_{10}$length to $\log_{10}$weight regression slopes and intercepts. Using ANCOVA we tested for effects of VI tagging on condition after accounting for effects of tributary drainage (13 drainages tested) and year of capture (within year tagged or 1 year post-tagging) on 4,960 observations. Statistical tests were made using the SAS Windows program (version 6.03; SAS Institute 1994) and Splus (version 3.1; Statistical Sciences 1993) with significance levels set to 0.05.

Results

We recaptured 348 previously tagged (clipped adipose fin) westslope cutthroat trout during the year they were marked. Of these, 201 (58%) had retained their tag. Of the 616 previously tagged westslope cutthroat trout recaptured the year after tagging, 355 (58%) had retained their tag. We observed a sigmoid relationship between tag retention and fish length at time of recapture (Figure 3).

Weighted logistic regression analyses indicated that fish length was the most significant variable which influenced tag retention (N=712; deviance improvement of 42.5 over intercept only model; P<0.001). Fish condition did not significantly improve the deviance performance of the length model (N=712; deviance improvement of 0.2 over intercept and length model; P=0.69).

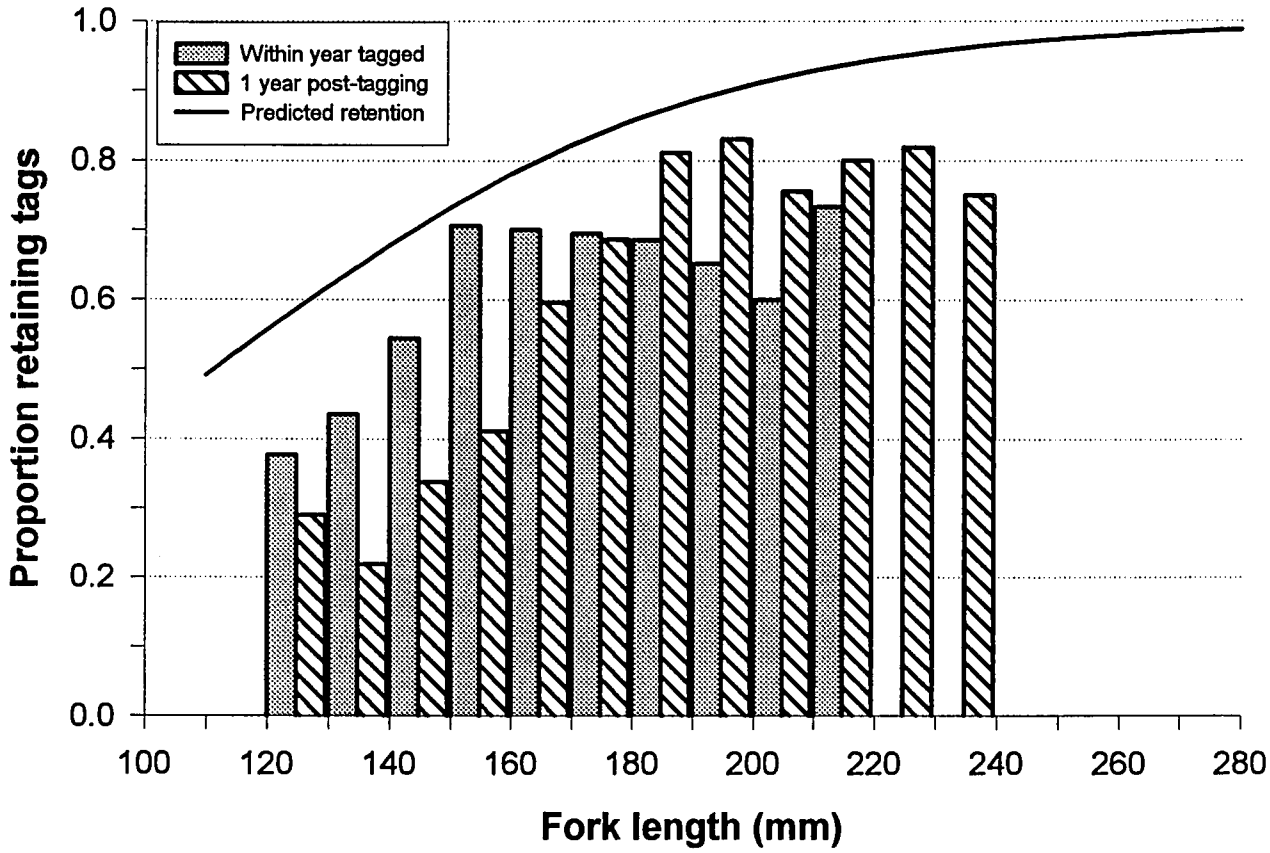


Figure 3. Proportional size distributions at recapture for westslope cutthroat trout that retained visible implant tags within the year of tagging (shaded bars) and 1 year after tagging (crosshatched bars), and the predicted retention of visible implant tags based on length at tagging (solid line).

Because condition factor did not add significantly to the model, we added the 58 fish that did not have weight measurements back into the sample. We then tested effects of tagging quality rating (R), so we removed the 59 fish for which tagging quality had not been rated from the sample. Fish length was still the most significant variable influencing tag retention ($P < 0.001$; Table 2). Tag rating, wetted stream width, and reach gradient significantly improved the model. Tag retention improved as fish length increased, as tag insertions improved, and as streams became narrower and of lower gradient. The model which used tributary drainage as a class variable was not significantly better than the model which used reach estimates of wetted width and channel gradient, suggesting that drainage effects were not as important as reach effects. There was not a significant difference ($P > 0.10$) in slopes or intercepts of $\log_{10}$length-$\log_{10}$weight regressions between those recaptured fish which had retained versus lost their tags.

After accounting for differences between drainages and year ($P < 0.01$; ANOVA), fish condition was not significantly different (ANCOVA; $P > 0.95$) between previously tagged and untagged westslope cutthroat trout. There was also not a significant difference in slopes ($P > 0.50$) or intercepts ($P > 0.05$) of $\log_{10}$length-$\log_{10}$weight regressions between fish which had previously been tagged and newly captured fish $\geq$ 110 mm across all drainages. However, when the 11 drainages which had adequate sample sizes were run individually, significant ($P < 0.05$) differences were found between $\log_{10}$length-$\log_{10}$weight regression slopes for previously tagged versus untagged fish in one of the drainages and between regression intercepts in an additional three drainages.

Ninety-five percent of all tags were easily readable at recapture (526 of 556) with 96% (193 of 201) readable at recapture within the year tagged and 94% (333 of 355) readable a year after tagging. Only one (< 1%) recaptured tag was unreadable within the same year as it was implanted, while three tags (1%) were unreadable one winter after tagging. Tags were unreadable either because they had been inserted too deeply into opaque tissue or because adipose tissue had clouded over the tags. We subsequently found that unreadable tags could be read if they were recoverable using the syringe.

Discussion

While the proportion of recovered westslope cutthroat trout which retained VI tags within the year tagged (58%; N=348) or the year following tagging (58%; N=616) was consistent, our estimated tag retention was lower than most previous studies of VI tag retention. These results can be partially explained by the small lengths at which we tagged the majority of our fish (1,145, or 55%, were <150 mm; FL). Using logistic regression we predicted tag retention (P_r) based on fork length (FL) and "good" tagging quality at time of tagging (Figure 3):

$$P_r=[e^{(-2.894+0.026*FL)}]/[1+e^{(-2.894+0.026*FL)}].$$

Table 2. Results of weighted logistic regression analyses to assess the probability of retention of visible implant tags in westslope cutthroat trout showing the improvement in model deviance the last entered explanatory variable provided. Explanatory variables are fork length at tagging (FL), tag insertion quality rating (R), wetted stream width (W), stream channel gradient as a percentage (G), and drainage (LOC). Explanatory variable coefficients and their associated standard errors (in parentheses) are shown for each model in variable order starting with the intercept, where the last explanatory variable added significantly to the model. Model run only for tagging data where tag insertions were rated.

Covariate(s)	df	Deviance improvement	P	Deviance	Coefficients (SE)
Intercept	499	598.5			
FL	498	551.1	47.4	<0.0001	-2.888(0.612) + 0.025(0.004)*FL
FL+R	496	544.0	7.1	0.029	0.026(0.004)*FL - 2.894(0.615), when tag rating "Good" - 3.406(0.676), when tag rating "Fair" - 5.212(1.494), when tag rating "Poor"
FL+R+W	495	512.8	31.2	<0.001	0.029(0.004)*FL - 1.873(0.659) , when tag rating "Good" - 2.348(0.722), when tag rating "Fair" - 4.056(1.403), when tag rating "Poor"
FL+R+W+G	494	496.4	16.4	<0.001	0.027(0.004)*FL + 0.611(0.908), when tag rating "Good" + 0.181(0.963), when tag rating "Fair" - 1.706(1.512), when tag rating "Poor" - 1.069(0.163)*W - 0.291(0.074)*G
FL+R+LOC	483	482.5	13.9	0.382	

Note that lengths for "within year tagged" and "1 year post-tagging" are lengths at recapture, while lengths for the prediction equation are lengths at tagging.

Our predicted retention rates were consistent with observed retention rates of 94% for 200-307 mm (FL) hatchery raised sea-run cutthroat trout O. clarki released into the Cowlitz River, Washington (Blankenship and Tipping 1993) and 82% retention estimated for rainbow trout 140-240 mm (TL)(Mourning et al. 1994). However, our observed retention rates were slightly higher than the 50% retention observed in brook trout 130-160 mm (TL) and slightly lower than the 100% retention observed for brook trout 200 mm and longer in the wild (Bryan and Ney 1994). Our predicted and observed tag retention rates indicated that the manufacturer's recommended minimum tagging length of 150 mm (TL) would result in at least a 73% retention rate for fish this length and longer. For retention rates over 90% the minimum size at tagging for westslope cutthroat trout would have to increase to about 195 mm.

Quality of tag insertion (tag rating) was significantly related to retention. This result was not unexpected because we assumed that "good" tag insertions resulted in better tag retention. This result emphasizes the need to have experienced personnel do the tagging.

Tag retention varied by drainage, but most drainage variation appeared to be explained by estimates of wetted width and channel gradient made in each reach. These variables are relatively easy to estimate. We are uncertain of the mechanism by which these two variables affect tag retention. We speculate that since higher channel gradients indicate faster water velocities, tag retention might be related to water velocity. Niva (1995) reported that VI tag retention was related to immediate post-tagging handling, with a higher proportion of fish dropped into the water losing tags than fish gently placed into the water. Release of VI tagged fish into high or highly variable velocities might cause similar tag loss. Differences in tag retention rates may also be related to differences between species or environments (McMahon et al. 1996; P. Bergman, Northwest Marine Technology, personal communication). We recommend further research to determine if, and why, different environments, or different fish stocks, lead to different tag retention rates.

We expected fish in better condition to retain tags at higher rates since we initially believed that adipose tissue at tag insertion sites was a form of fatty, or excess, tissue. However, fish condition did not significantly improve deviance performance in the logistic regression model nor were $\log_{10}$length-$\log_{10}$weight regression slopes or elevations significantly different between recaptured fish that retained versus lost tags. Our inability to find a significant relationship between fish condition and tag retention led us to question or original assumption that post-orbital adipose tissue is fatty, or excess, tissue. We have subsequently discovered that the post-orbital, clear tissue at tag insertion sites is primarily a stroma (matrix) of extremely fine, microfibrils of collagen, a form of connective tissue. This tissue contains a few fibrocytes (which form and maintain the collagen), blood vessels, and sinuses (J. Morrison, Fish and Wildlife Service, personal communication).

We found no significant difference ($P > 0.95$; ANCOVA) in fish condition between previously tagged and untagged westslope cutthroat trout in our study. This finding is consistent with

findings of Bryan and Ney (1994) who found no significant differences ($P > 0.2$) between condition factors of VI tagged and untagged brook trout in the wild. We also found no significant differences between $\log_{10}$length-$\log_{10}$weight regression slopes ($P > 0.50$) and regression elevations ($P > 0.05$) for previously tagged versus untagged fish. However, we did observe significant differences ($P < 0.05$) in either slopes, or elevations for non-significant slope differences, for four of 11 drainages tested. We are uncertain if those differences in fish condition we observed were caused by tagging effects or factors associated with electrofishing and handling. Since we do not have good evidence that VI tags affected condition of tagged fish, our evidence suggests that VI tags do not affect growth or, inferentially, survival. We recommend that studies be conducted to further test potential effects of VI tags on growth and survival of tagged fish.

Since we used adipose clips as a secondary mark, our efforts to find VI tags on recaptured fin-clipped trout were probably more thorough than if no secondary mark had been used. Since Coombs et al. (1990) found little (0.2%) adipose fin regeneration, or effect of fin removal on growth or survival on Atlantic salmon 3 months following fin clipping, we believe that this mark identified all tagged fish without affecting their growth or survival. Use of this secondary mark might have led to a slight positive bias in our assessment of tag retention.

We conclude that VI tags provide a valuable means to individually mark small fish that typically make up headwater populations. However, investigators must recognize that a relatively high rate of tag loss (> 25%) may occur in fish under 150 mm because tag retention is strongly and positively related to fish length. In spite of the relatively low tag retention rates in these smaller fish, we obtained valuable information on movements for fish that retained their tags. Tag insertion quality positively influenced tag retention and should be considered and documented in any tagging study.

When using VI tags we recommend: 1) personnel inserting the tags must be experienced (Niva 1995) and records should be kept of tag insertion quality; 2) tagging and tag recovery protocols must be established and followed (Geoghegan et al. 1990); 3) if it is necessary to use different tag colors, select colors which contrast sharply to enhance ability to recognize colors upon recapture (Niva 1995); and 4) tag retention should be evaluated (Vreeland 1990; McFarlane et al. 1990; Nielsen 1992). We developed a predictive equation to estimate tag retention by length at tagging for westslope cutthroat trout 100 to 300 mm (FL) and suggest that future investigators could adopt this protocol to evaluate tag retention.

Chapter 2:
Movement of resident westslope cutthroat trout inhabiting isolated headwater streams of Montana

Bradley B. Shepard
Montana Department of Fish, Wildlife and Parks
Department of Biology
Montana State University, Bozeman, Montana 59717, USA

and

Susan C. Ireland and Robert G. White
Montana Cooperative Fisheries Research Unit[a]
Department of Biology, Montana State University
Bozeman, Montana 59717, USA

[a] The Unit is jointly sponsored by Montana State University, Montana Department of Fish, Wildlife, and Parks, and the U.S. National Biological Service.

Introduction

Home range has been defined as "the area over which the animal normally travels" (Hayne 1949; Gerking 1953). Home range differs from territory in that territories are usually relatively smaller areas that are actively defended (Noble 1939; Nice 1941; Gerking 1953), while home ranges are usually larger areas that are not defended. Brook trout (Shetter 1937), brown trout (Schuck 1943), cutthroat trout (Miller 1957), Atlantic salmon (Saunders and Gee 1964), and other species (Gerking 1959) have been reported to remain in a relatively restricted home range during some or all of their life phases. Gerking (1953) suggested that for fishes that inhabit streams containing riffle-pool structure "each pool can be considered as a more or less isolated unit containing a natural population of its own". Gerking (1953) further hypothesized that "the fish population of a small stream may be thought of as a series of discrete, natural units rather than as a single, homogeneous, freely-mixing group". Allen's (1951) study of brown trout in the Horokiwi Stream of New Zealand and Miller's (1957) study of cutthroat trout in a small Colorado stream support these views. Gerking (1953) made it clear that while restricted home ranges may exist for most individuals, some individuals may not remain within restricted home areas. He emphasized that stray fish were very important for potentially occupying territories left vacant and spreading the species over the geographic limits of its tolerance.

Rieman and McIntyre (1993) discussed the role of movement and dispersal in population structuring of bull trout. We investigated the scale of movement that occurred in populations of westslope cutthroat trout that inhabited headwater portions of tributary streams. Since most of these populations were isolated, we also investigated the relationship between the level of isolation and the amount of movement we documented for each population.

Study Area

Movement of westslope cutthroat trout was assessed in 13 tributary drainages of the upper Missouri River and two tributary drainages of the upper Clark Fork River, Montana (Figure 1). Streams in these tributary drainages were characterized by low-flow wetted widths of 1.4 to 6.6 m, water conductivities of 64 to 661 umhos, elevations of 1320 to 3000 m, and gradients of 3 to 11% (Table 1). Cutthroat trout in sampled populations exhibited little evidence of introgression with either rainbow trout Oncorhynchus mykiss or Yellowstone cutthroat trout Oncorhynchus clarki bouveri using external morphometric characteristics and/or electrophoretic testing (Salmon and Trout Genetics Laboratory, University of Montana, Missoula, Montana, unpublished data).

Methods

Fish were captured using a generator powered Smith-Root BP-15 backpack electrofisher ™ (product names are for clarification of methods and do not represent product endorsement). We used frequencies under 40 Hz at pulse widths of at least 1 msec and voltages between 100 to 700 volts to maximize the number of fish captured, while minimizing injury to the fish caused by the shock (W. Fredenberg, Fish and Wildlife Service, personal communication). All captured fish were measured to the nearest mm (FL) and weighed to the nearest gram using either a Ohaus LS2000 or C505 battery powered scale. Most Oncorhynchus spp. over 120 mm were tagged with individually numbered visible implant (VI) tags (available from Northwest Marine Technology, Shaw Island, Washington) using a syringe (Haw et al. 1990; Kincaid and Calkins 1992). Some cutthroat over 120 mm were not tagged because either: 1) their post-orbital adipose tissue was so fragile a VI tag would not stay in place; or 2) they were worked after dark tags could not be read; or 3) when the tagging syringe froze solid during late fall sampling.

We measured the length of stream channels from a fixed point (usually the stream's mouth) on 1:24,000 USGS maps with a map wheel to establish reference points for each sample section. Reference distances were assigned to the middle of each sample section and these stream kilometer distances were used to compute movement distances. We verified our map measurements by measuring stream lengths in the field using a hip chain over at least a 1.5 km distance in three streams. Map measurements were within 0.1 km of field measurements. Lengths of sample sections were generally 50 to 150 m in length. We believe our estimates of movement were accurate to about 100 m.

Gowen et al. (1994) and Gowen and Fausch (1996) cautioned that traditional methods for assessing movement may lead to erroneous conclusions and made some suggestions regarding

study design to document movement. Movement was assessed from recaptures of VI tags. VI tags were implanted in Oncorhynchus spp. captured during all sampling events from 1992 through 1994. Recoveries of tags occurred from 1993 through 1995. Most sampling occurred during the summer period from late-June through mid-September. However, from late-September through early November of 1993 and from May through June of 1994 we re-sampled all permanent sample sections as well as sections adjacent to all sample sections in an effort to recover previously marked fish and to tag fish over a larger proportion of the stream. Shepard et al. (1996; Chapter 1 this report) evaluated factors affecting retention of visible implant tags in westslope cutthroat trout. They found tag retention to be strongly related to fish length at tagging and retention rates averaged slightly over 50%.

Movement of recaptured tags were plotted as histograms showing both the direction of movement and absolute value of distance moved for those drainages where at least 20 recapture events had occurred. In addition, Spearman rank correlation coefficients were computed between the proportion of recapture events within a drainage where a fish had moved at least 0.5 km and the rated level that the habitats supporting that fish population were isolated by some type of barrier to fish movement upstream (Table 3).

Table 3. Ranks assigned to levels of isolation for westslope cutthroat trout populations where movement was evaluated from 1993 to 1995.

Rank	Isolating mechanism
0	No known isolation.
1	Isolated from downstream habitats due to recent (anthropgenic – culverts, dams, etc.) causes.
2	Isolated from downstream habitats during all times except during extreme high stream flow events due to natural intermittent portions.
3	Isolated from downstream habitats due to a geologic barrier such as waterfall.

Results

We sampled relatively high proportions of the habitat occupied by westslope cutthroat trout in sample drainages (Table 4). For most streams we sampled at least 10% of the occupied habitat during at least one sampling occasion and for some streams sampling occurred over more than 50% of occupied habitats during at least one sampling event.

A large proportion (580 of 841 recaptures; 70%) of tag recaptured fish had remained within the sample section in which they had been tagged or last recaptured (Figure 4). We observed some slight differences in the numbers of fish that had moved between stream drainages (Figure 5). The proportion of tag recaptured fish that had moved over 0.5 km was significantly and inversely correlated to the ranked level of isolation (Spearman correlation coefficient = -0.621; $P<0.05$).

Table 4. Length (km) of habitat occupied by westslope cutthroat trout and length (%) of occupied habitats sampled by season and year (Su93 = Summer 1993, Fa93 = Fall 1993, Sp93 = Spring 1993, Su94 = Summer 1995, and Su95 = Summer 1995) by stream.

			Length sampled (% of occupied)				
Drainage	Stream	Occupied	Su 93	Fa 93	Sp 94	Su 94	Su 95
Collar Creek (Box Elder Creek)	Collar	2.70	2.09 (77)	0.53 (20)	0.54 (20)	0.64 (24)	0.47 (17)
Cottonwood (Ruby River)	Geyser	1.65	0.71 (43)	0.89 (54)	0.50 (30)	0.40 (24)	1.16 (70)
	Cottonwood	5.80	0.12 (2)	0.11 (2)	0.11 (2)	0.43 (7)	-
Cottonwood (Smith River)	Main Cottonwood	0.60	0.07 (12)	0.14 (23)	0.13 (22)	0.16 (27)	0.14 (23)
	E Fk Cottonwood	3.30	0.18 (5)	0.12 (4)	0.03 (1)	0.35 (11)	0.12 (4)
	W Fk Cottonwood	4.25	0.42 (10)	0.09 (2)	0.32 (8)	0.71 (17)	0.22 (5)
Douglas (Flint Creek)	N Fk Douglas	3.10	1.00 (32)	1.53 (49)	0.47 (15)	0.48 (15)	0.64 (21)
	Main Douglas	1.10	0.27 (25)	0.46 (42)	0.18 (16)	0.08 (7)	-
Gold (Clark Fork)	N Fk Gold	4.20	0.32 (8)	0.77 (18)	0.32 (8)	0.34 (8)	0.34 (8)
Halfmoon (Flatwillow Creek)	Halfmoon	7.30	0.35 (5)	-	-	0.70 (10)	0.67 (9)
Halfway (Jefferson River)	Halfway	9.60	0.61 (6)	0.95 (10)	0.56 (6)	0.74 (8)	0.20 (2)
Jerry (Big Hole River)	Delano	1.90	0.23 (12)	0.84 (44)	0.34 (18)	0.22 (12)	0.28 (15)
	Jerry	5.80	0.36 (6)	0.34 (6)	0.27 (5)	0.25 (4)	0.20 (3)
	Flume	1.00	0.03 (3)	0.05 (5)	-	-	-
McVey (Big Hole River)	McVey	3.30	2.36 (72)	0.75 (23)	1.14 (35)	0.48 (15)	0.45 (14)
Soap (W Fk Madison R)	Soap	3.90	0.40 (10)	0.95 (24)	0.27 (7)	0.40 (10)	0.42 (11)
Tenderfoot (Smith River)	Tenderfoot	8.1	0.76 (9)	0.76 (9)	-	0.97 (12)	0.85 (10)
White's (Missouri River)	White's	4.60	2.96 (64)	1.14 (25)	0.13 (3)	1.06 (23)	2.48 (54)

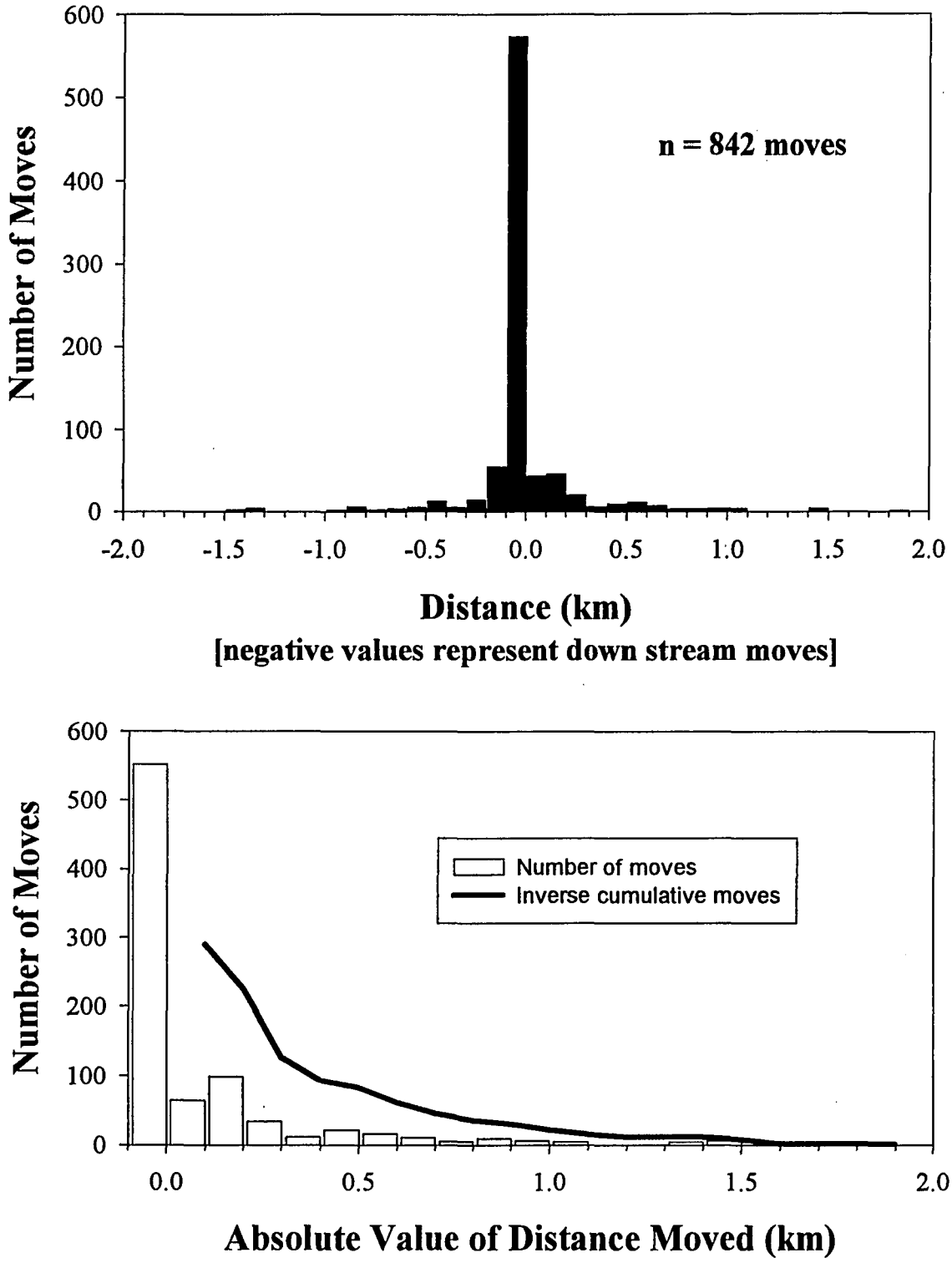


Figure 4. Distance that tagged and recaptured westslope cutthroat trout had moved upstream or downstream (upper graph) and the absolute value of the distance moved (lower graph) in 14 tributary drainages from 1993 to 1995.

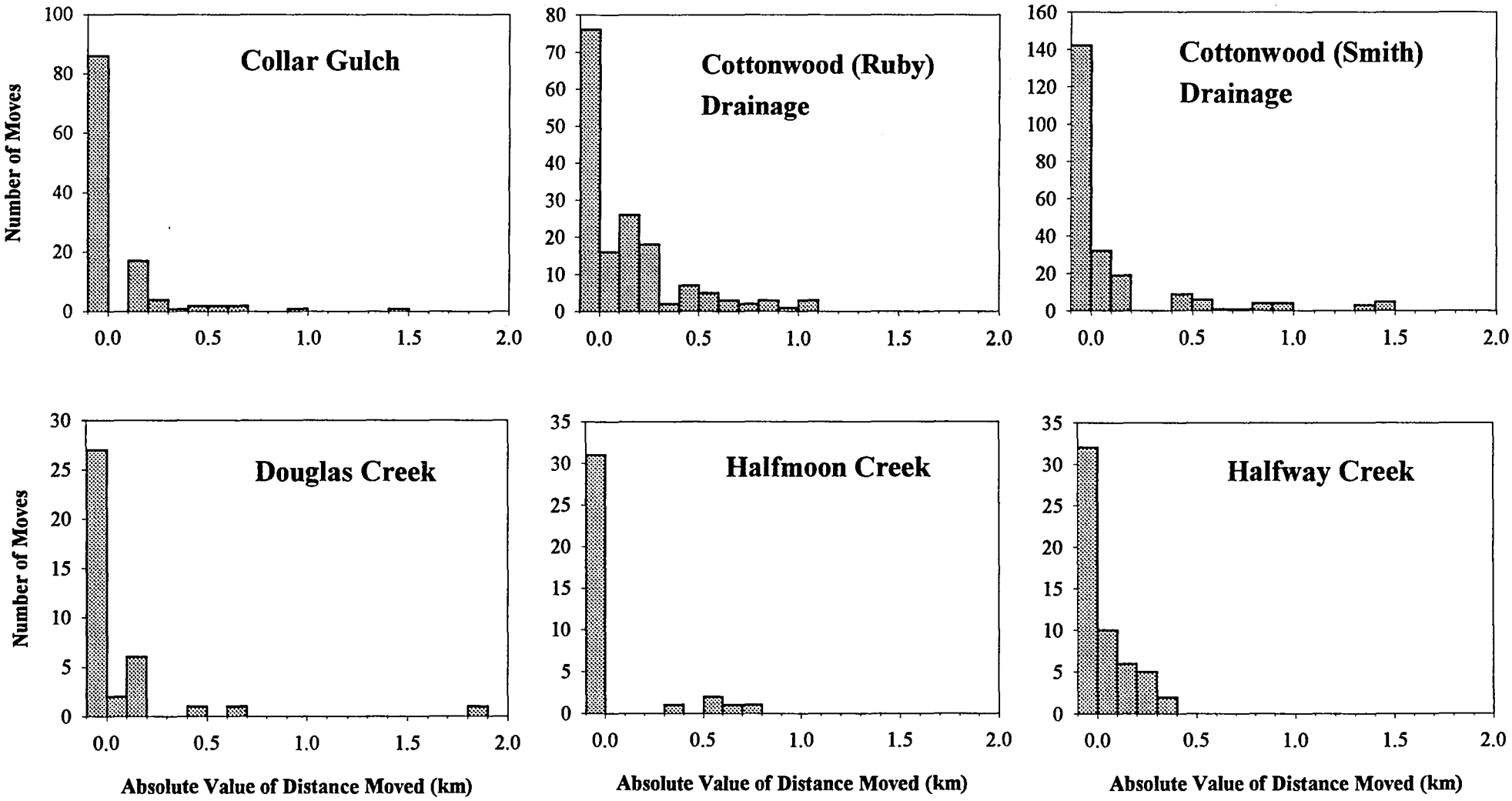


Figure 5. Absolute value of the distance that tagged and recaptured westslope cutthroat trout had moved in the Collar, Cottonwood (Smith), Cottonwood (Ruby), and Douglas, Halfmoon, and Halfway drainages during the period 1993 to 1995.

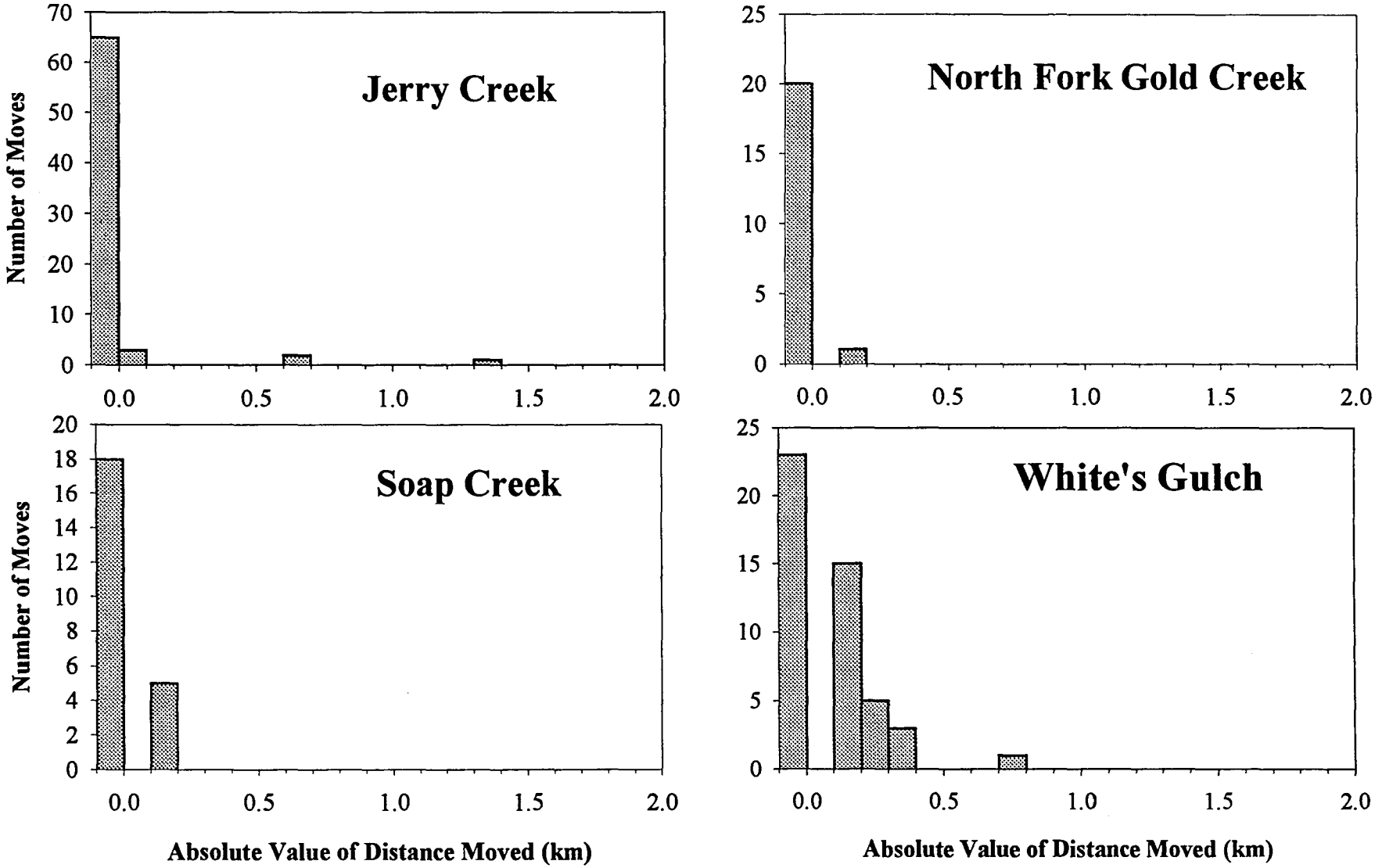


Figure 6. Absolute value of the distance that tagged and recaptured westslope cutthroat trout had moved in the Jerry, North Fork Gold, Soap, and White's drainages during the period 1993 to 1995.

Discussion

We sampled relatively large proportions of occupied habitats (Table 4). This extensive sampling should have reduced movement bias associated with small sample size. Our recapture information suggests that relatively little movement is occurring in these resident westslope cutthroat trout populations, but that a few individuals may be moving relatively long distances. Since we did not set downstream traps at the lower boundaries of population distributions nor sample much below the known distribution of westslope cutthroat trout, fish that moved downstream out of the habitats that supported the majority of the population would not have been sampled. We documented that a few large fish moved downstream into segments of stream channels in both Collar and Cottonwood (Smith drainage) creeks that only carried stream flow during high flow events. We also documented that large fish occupying a sample section in Halfway Creek apparently left that section and moved out of the headwater reaches because we did not recapture these larger fish after tagging them. This sample section was located below an old mining settling pond and large fish appeared to move down out of this pond and then were unable to move back up into the pond. The proportion of documented movements 0.5 km or longer was correlated with the level of isolation experienced by the fish population.

Chapter 3:
Length frequencies of westslope cutthroat trout resident of headwater streams in Montana

Bradley B. Shepard and Robert G. White
Montana Cooperative Fishery Research Unit[a]
Biology Department
Montana State University
Bozeman, Montana 59717-0001

[a] The Unit is jointly sponsored by Montana State University, Montana Department of Fish, Wildlife, and Parks, and the U.S. National Biological Service.

Introduction

This chapter presents summarized length-frequency data for westslope cutthroat trout sampled from 1993 through 1995 in headwater streams of Montana (Figure 1). Fork lengths were measured to the nearest millimeter. Sampling was done primarily during the summer, from June 20 through September 15. Additional sampling was done during fall 1993 from September 22 through November 4, and during spring 1994 from May 24 to June 16.

Results and Discussion

The length-frequency histogram for all westslope cutthroat trout sampled from all sites during the summer season showed two nodes that could possibly be used for separating age classes (Figure 7). A node appeared from 60 to 90 mm that probably indicated age 1 fish and another slight node existed between 100 to 140 that might indicate age 2 fish. Age 0 fish were generally under 50 mm and in most streams age 0 fish had not emerged from the streambed until late summer or early fall. In addition, age 0 fish were not always captured because they slipped through dip nets due to their extremely small size and were so fragile that we chose not to capture, hold, and work them during most sampling events.

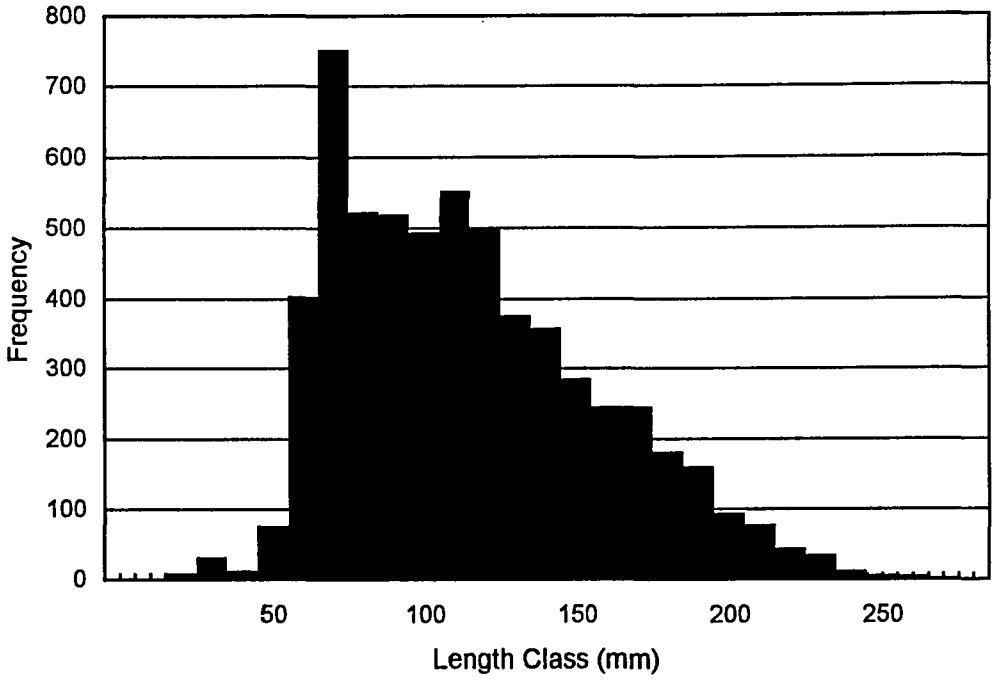


Figure 7. Length-frequency histogram for the summer season (June 20 to September 15) combined for all years (1993 through 1995) and for all streams.

Length-frequency histograms were constructed for individual streams where at least 100 westslope cutthroat trout had been sampled during at least one of the three years (Figures 8 to 10). These length-frequency histograms showed relatively clear nodes for age 1 and 2 fish for some streams. Age 3 fish could also sometimes be identified using these histograms for a few streams. The histogram for the West Fork of Cottonwood Creek shows relatively strong nodes for ages 1, 2, and 3 (Figure 10). However, ages were not easily interpreted from all the histograms.

For some streams the histograms showed that a year class had been severely depressed. For instance, it appeared that very few age 1 fish (60 to 90 mm) were present in Collar Creek during 1993 (the 1992 year-class). This missing year-class was also obvious in 1994 (low numbers of fish between 100 and 140 mm) and 1995 (low numbers between 140 and 160 mm; Figure 8). Delano, Geyser, Soap, and West Fork Cottonwood creeks also appeared to have weak 1992 year-classes; however, Soap Creek was the only stream where this poor year-class could be easily seen in histograms from subsequent years (Figures 8 and 10). Conversely, the remaining streams seemed to have supported a good 1992 year-class based on 1993 histograms.

We suspect that growth that occurred throughout the summer months may have made the interpretation of ages from summer length-frequency histograms difficult. We had hoped that histograms for fish collected during the fall and spring might have been easier to interpret, but we did not find that to be the case (Figures 11 and 12). We also concur with Downs' (1995) suggestion that different fish inhabiting different microhabitats within a single stream

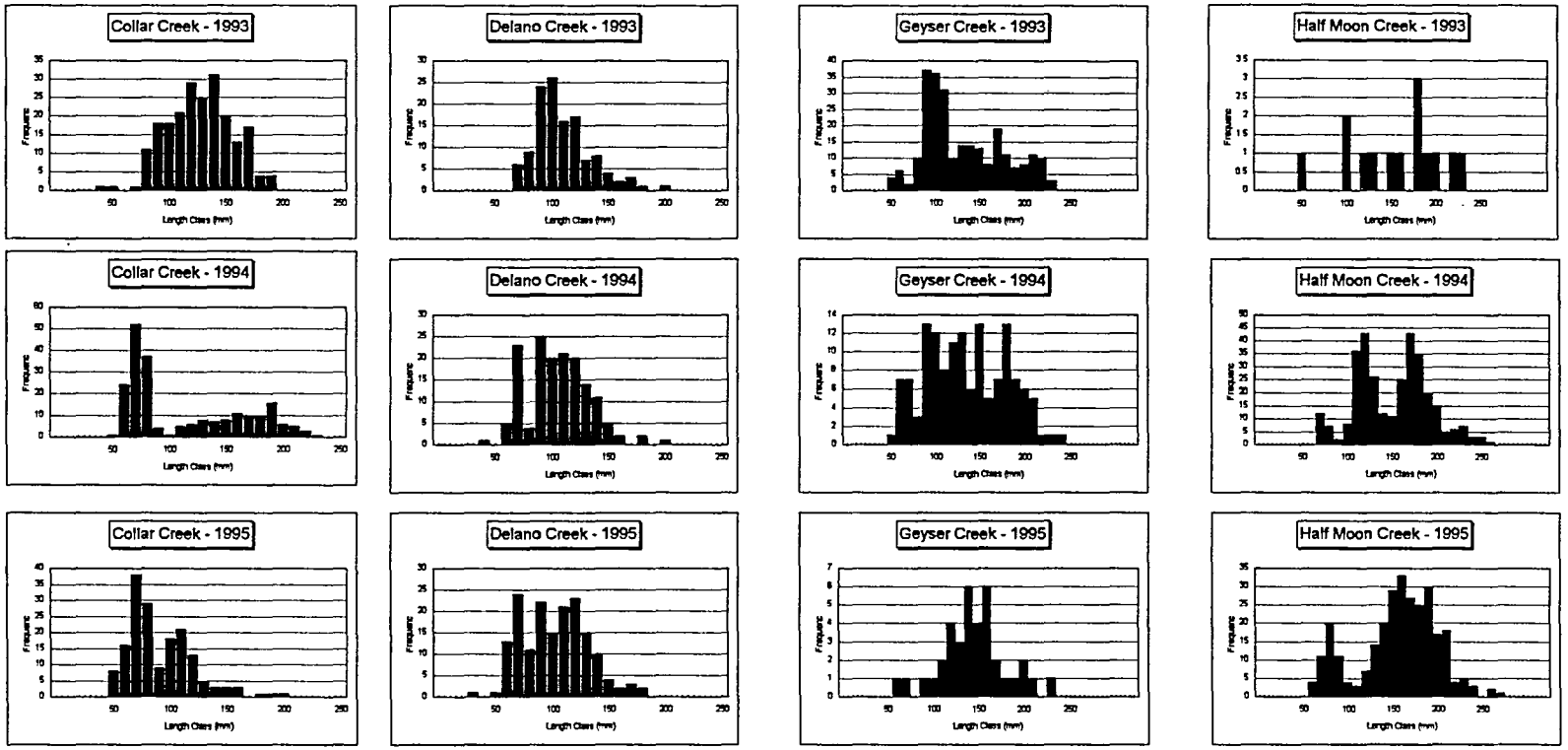

Collar Creek - 1993
Delano Creek - 1993
Geyser Creek - 1993
Half Moon Creek - 1993
Collar Creek - 1994
Delano Creek - 1994
Geyser Creek - 1994
Half Moon Creek - 1994
Collar Creek - 1995
Delano Creek - 1995
Geyser Creek - 1995
Half Moon Creek - 1995

Figure 8. Length-frequency histograms for westslope cutthroat trout sampled from Collar, Delano, Geyser, and Halfmoon creeks during the summers of 1993 through 1995.

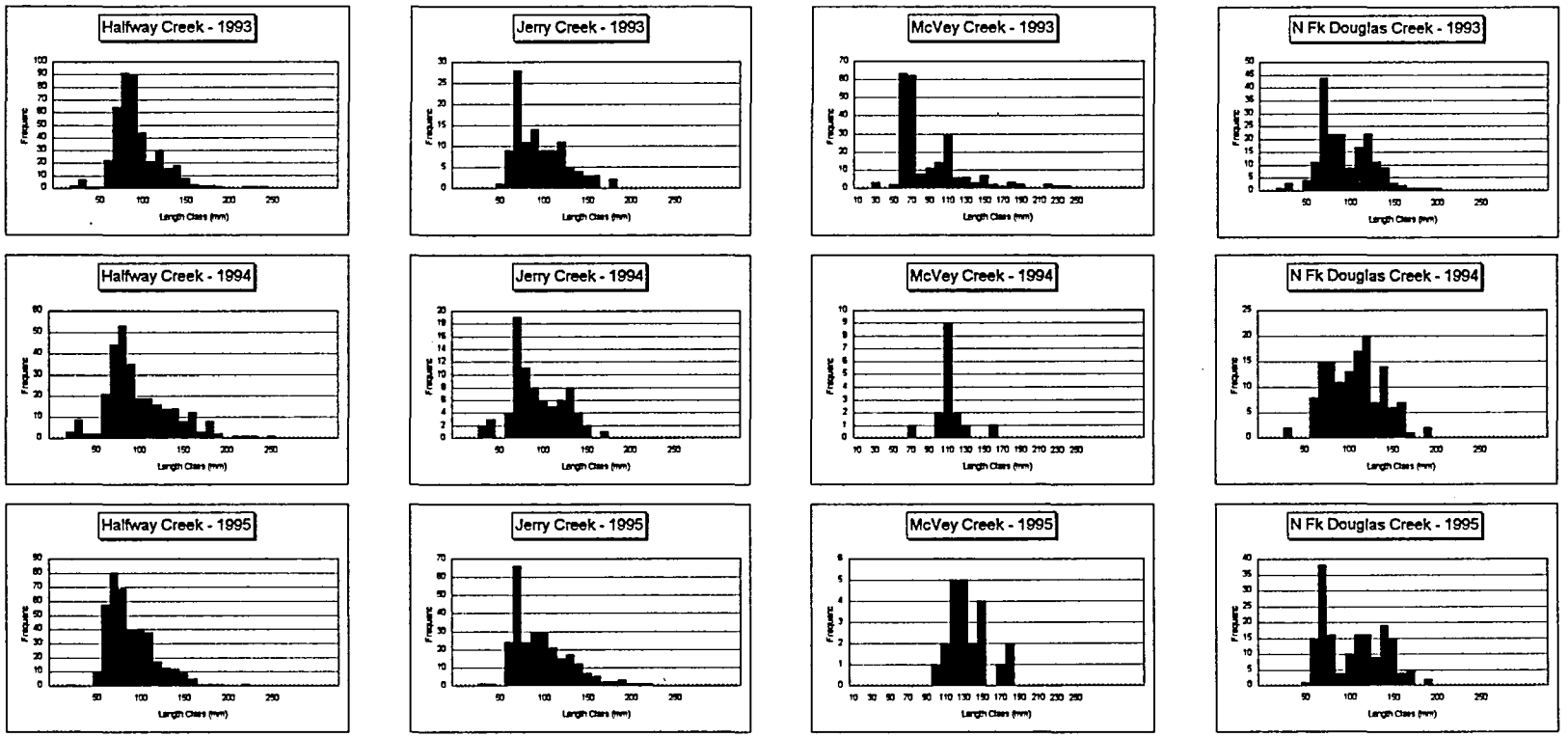

Halfway Creek - 1993
Jerry Creek - 1993
McVey Creek - 1993
N Fk Douglas Creek - 1993
Halfway Creek - 1994
Jerry Creek - 1994
McVey Creek - 1994
N Fk Douglas Creek - 1994
Halfway Creek - 1995
Jerry Creek - 1995
McVey Creek - 1995
N Fk Douglas Creek - 1995

Figure 9. Length-frequency histograms for westslope cutthroat trout sampled from Halfway, Jerry, McVey and North Fork Douglas creeks during the summers of 1993 through 1995.

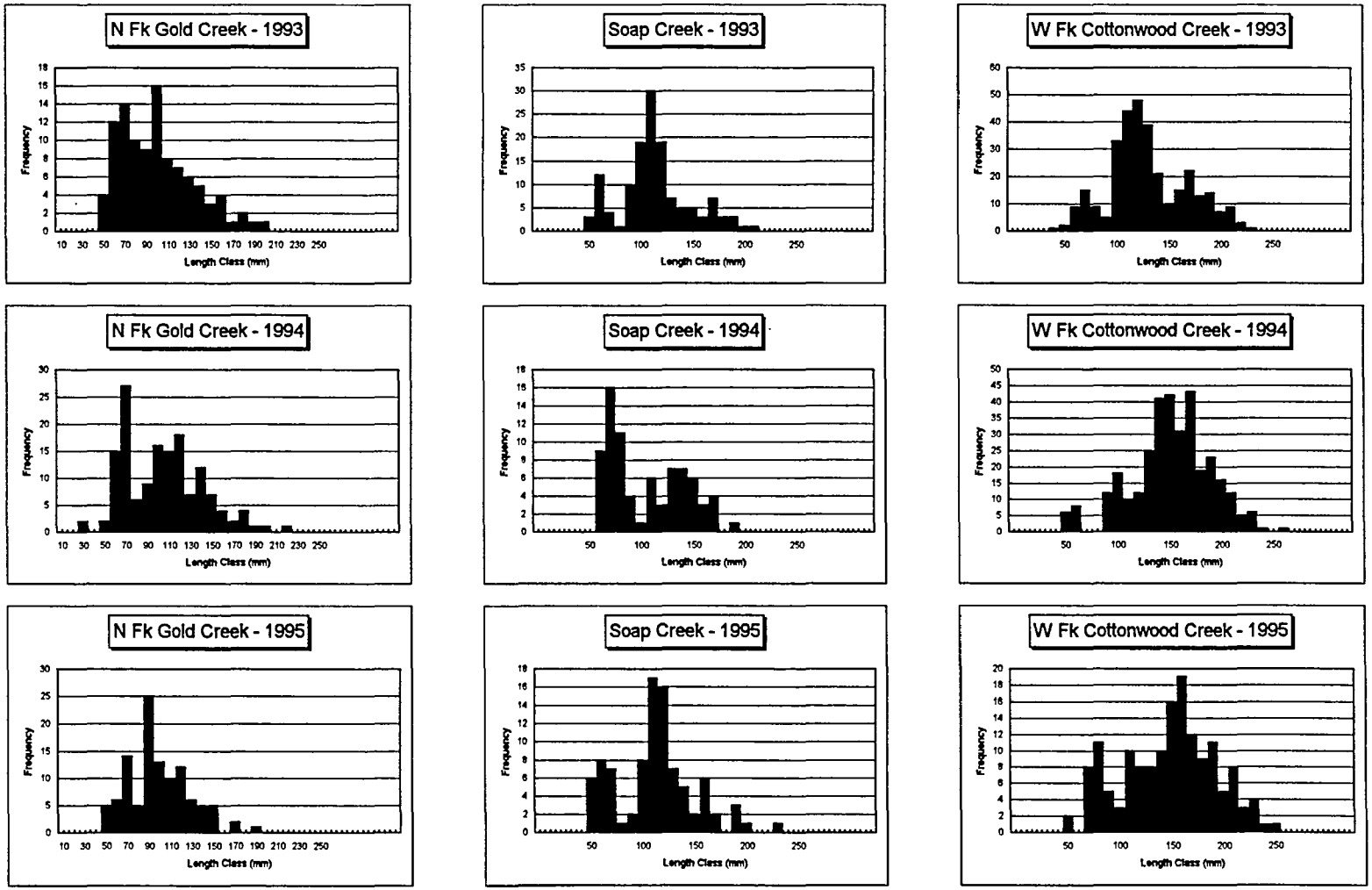

N Fk Gold Creek - 1993
Soap Creek - 1993
W Fk Cottonwood Creek - 1993
N Fk Gold Creek - 1994
Soap Creek - 1994
W Fk Cottonwood Creek - 1994
N Fk Gold Creek - 1995
Soap Creek - 1995
W Fk Cottonwood Creek - 1995
Frequency
Length Class (mm)

Figure 10. Length-frequency histograms for westslope cutthroat trout sampled from North Fork Gold, Soap, and West Fork Cottonwood creeks during the summers of 1993 through 1995.

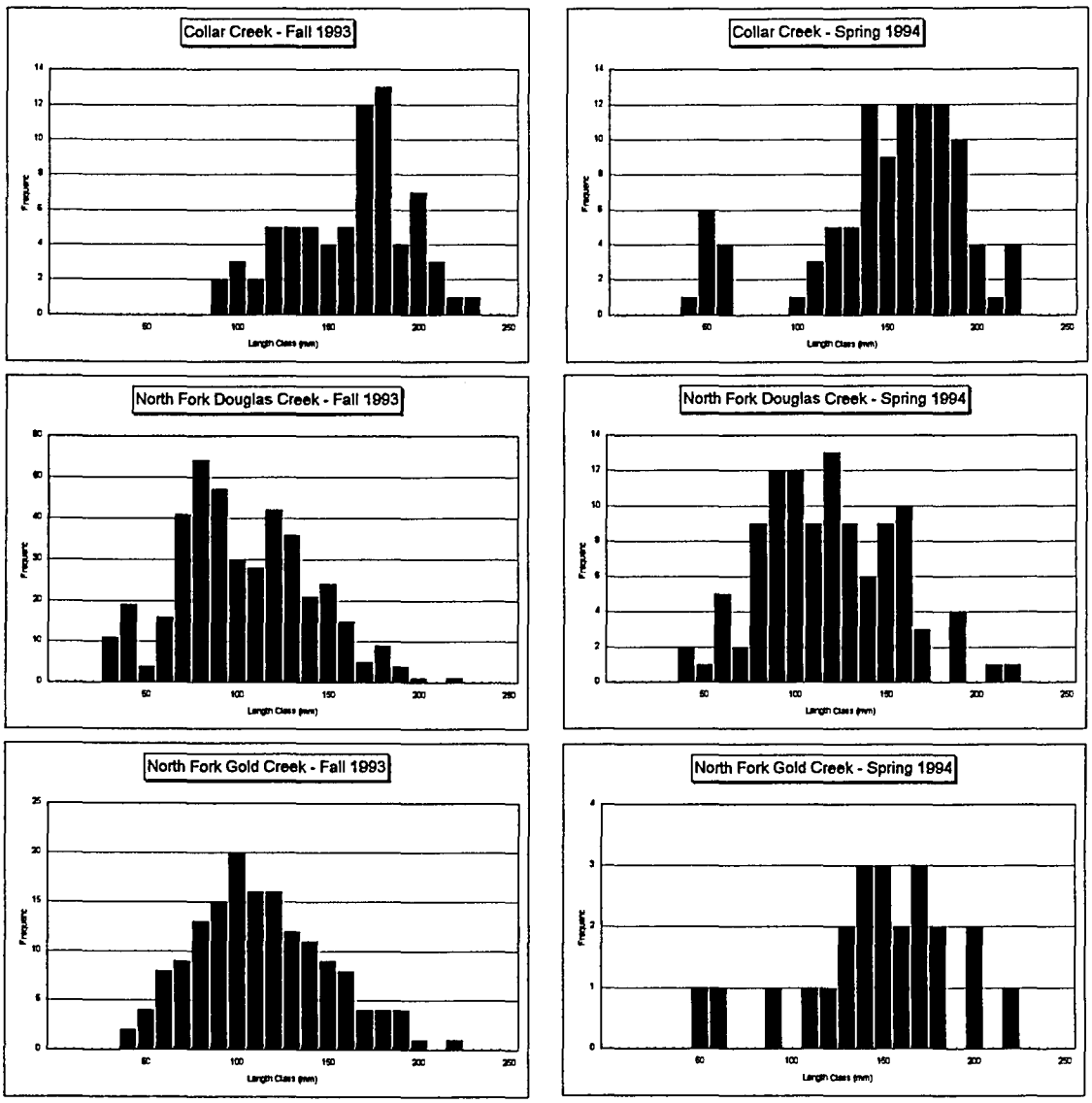

Collar Creek - Fall 1993
Collar Creek - Spring 1994
North Fork Douglas Creek - Fall 1993
North Fork Douglas Creek - Spring 1994
North Fork Gold Creek - Fall 1993
North Fork Gold Creek - Spring 1994

Figure 11. Length-frequency histograms for westslope cutthroat trout sampled from Collar, North Fork Douglas, and North Fork Gold creek, during the fall of 1993 and spring of 1994.

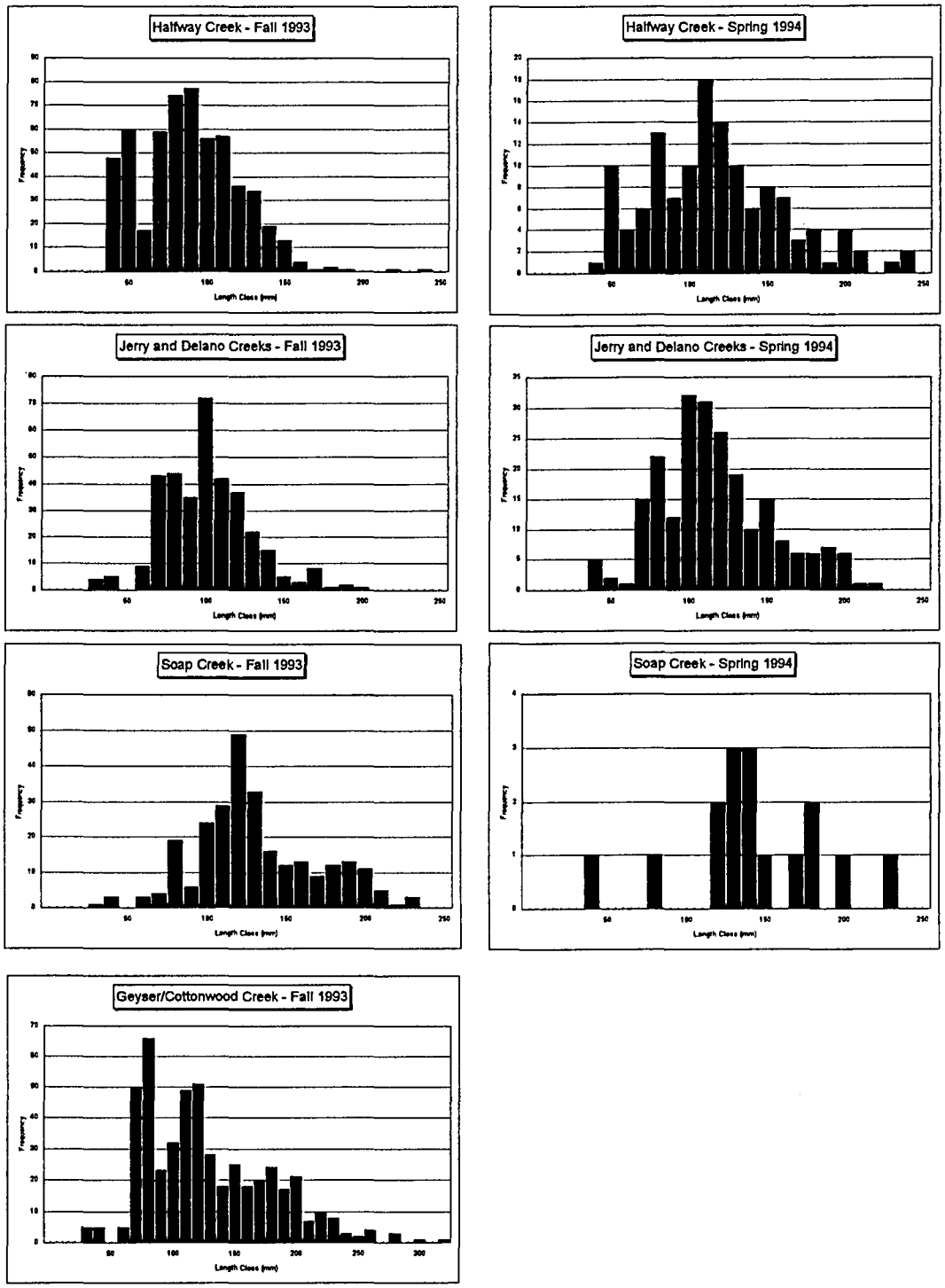

Halfway Creek - Fall 1993
Halfway Creek - Spring 1994
Jerry and Delano Creeks - Fall 1993
Jerry and Delano Creeks - Spring 1994
Soap Creek - Fall 1993
Soap Creek - Spring 1994
Geyser/Cottonwood Creek - Fall 1993

Figure 12. Length-frequency histograms for westslope cutthroat trout sampled from Halfway, Jerry/Delano, Soap, and Geyser/Cottonwood creeks during the fall of 1993 and spring of 1994.

may exhibit different growth rates. If true, this mechanism would also make it more difficult to interpret ages using length-frequency histograms.

In conclusion, length-frequency histograms could not be easily used for assigning ages and assessing relative year-class strengths across all sampled streams. We believe that length-frequency histograms could be used for determining that a year-class was nearly absent, especially if the missing year-class was age 1 at the time length-frequency sampling was done. We recommend that any length-frequency sampling done to compare among streams or years be done at similar times of year and suggest using either early or late summer as a preferred sampling time. We further recommend that a minimum of 100 fish be measured from each stream at each sampling period to construct length-frequency histograms for size structure analyses. These fish should be collected over the length of habitat occupied by the population to ensure that all age classes that are present are sampled.

Acknowledgements

This research was supported by funds provided by the U.S. Department of Agriculture, Forest Service, Rocky Mountain (formerly Intermountain) Research Station; U.S. Department of Agriculture, Forest Service, Region One; Montana Department of Fish, Wildlife and Parks; Montana Cooperative Fishery Research Unit; Montana State University; and the U.S. Department of Interior, Bureau of Land Management's Headwaters and Judith Resource Areas. C. Downs, D. Nixdorf, J. Pravochek, and D. Fuller assisted with field data collection. Montana Department of Fish, Wildlife and Parks fisheries biologists provided data and R. Spoon, D. Oswald, and J. Brammer and their staffs assisted with field work. Forest Service personnel on the Beaverhead, Deerlodge, Helena, and Lewis and Clark Forests and Bureau of Land Management personnel from the Judith and Headwaters Resource Areas of the Bureau of Land Management provided data and field support. D. Dunlap, B. Peresinni, and D. Webb assisted with data analysis for Chapter 1. L. Ahlm, P. Bergman, L. Blankenship, S. Martin, T. McMahon, B. Rieman, and A. Zale reviewed earlier drafts of Chapter 1 and provided valuable comments.

References

Allen, K.R. 1951. The Horokiwi Stream. A study of a trout population. New Zealand Marine Department of Fisheries Bulletin 10:1-231.

Anderson, R.O., and S.J. Gutreuter. 1983. Length, weight, and associated structural indices. Pages 283-300 in L.A. Nielsen, and D.L. Johnson, editors. Fisheries Techniques, American Fisheries Society, Bethesda, Maryland.

Blankenship, H.L., and J.M. Tipping. 1993. Evaluation of visible implant and sequentially coded wire tags in sea-run cutthroat trout. North American Journal of Fisheries Management 13:391-394.

Bryan, R.D., and J.J. Ney. 1994. Visible implant tag retention by and effects on condition of a stream population of brook trout. North American Journal of Fisheries Management 14:216-219.

Coombs, K.A., J.K. Bailey, C.M. Herbinger, and G.W. Friars. 1990. Evaluation of various external marking techniques for Atlantic salmon. American Fisheries Society Symposium 7:142-146.

Downs, C.C. 1995. Age determination, growth, fecundity, age at sexual maturity, and longevity for isolated, headwater populations of westslope cutthroat trout. Master's thesis. Montana State University, Bozeman, Montana.

Geoghegan, P., M.T. Mattson, D.J. Dunning, and Q.E. Ross. 1990. Improved data in a tagging program through quality assurance and quality control. American Fisheries Society Symposium 7:714-719.

Gerking, S. D. 1953. Evidence for the concepts of home range and territory in stream fishes. Ecology 34:347-365.

Gerking, S.D. 1959. The restricted movement of fish popuations. Biology Review 34:221-242.

Gowan, C. and K. D. Fausch. 1996. Mobile brook trout in two high-elevation Colorado streams: re-evaluating the concept of restricted movement. Canadian Journal of Fisheries and Aquatic Sciences 53:1370-1381.

Gowan, C., M. K. Young, and K. D. Fausch. 1994. Restricted movement in resident stream salmonids: a paradigm lost? Canadian Journal of Fisheries and Aquatic Sciences 51:2626-2637.

Haw, F., P.K. Bergman, R.D. Fralick, R.M. Buckley, and H.L. Blankenship. 1990. Visible implanted fish tag. American Fisheries Society Symposium 7:311-315.

Hayne, D.W. 1949. Calculation of size of home range. Journal of Mammology 30:1-18.

Hosmer, D.W. and S. Lemeshow. 1989. Applied logistic regression. John Wiley and Sons, New York, New York.

Kincaid, H.L., and G.T. Calkins. 1992. Retention of visible implant tags in lake trout and Atlantic salmon. Progressive Fish-Culturist 54:163-170.

McFarlane, G.A., R.S. Wydoski, and E.D. Prince. 1990. Historical review of the development of external tags and marks. American Fisheries Society Symposium 7:9-29.

McMahon, T.E., S.R. Dalbey, S.C. Ireland, J.P. Magee, and P.A. Byroth. 1996. Field evaluation of visible implant tag retention in brook trout, cutthroat trout, rainbow trout, and arctic grayling. North American Journal of Fisheries Management 16:.921-925

Miller, R.B. 1957. Permanence and size of home territory in stream-dwelling resident trout. Journal of the Fisheries Research Board of Canada 14:687-691.

Mourning, T.E., K.D. Fausch, and C. Gowan. 1994. Comparison of visible implant tags and Floy anchor tags on hatchery rainbow trout. North American Journal of Fisheries Management 14:636-642.

Nice, M.M. 1941. The role of territory in bird life. American Midlands Naturualist 26:441-487.

Nielsen, L.A. 1992. Methods of marking fish and shellfish. American Fisheries Society Special Publication 23, Bethesda, Maryland.

Niva, T. 1995. Retention of visible implant tags by juvenile brown trout, Salmo trutta. Journal of Fish Biology 46:997-1002.

Noble, G.K. 1939. The role of dominance in the life of birds. Auk 56:253-273.

Rieman, B. E. and J. D. McIntyre. 1993. Demographic and Habitat Requirements for Conservation of Bull Trout. General Technical Report INT-302, USDA Forest Service, Intermountain Research Station, Boise, Idaho.

Robison-Cox, J.R. 1998. A record linkage approach to imputation of missing data: analyzing tag retention in a tag-recapture experiment. Journal of Agriculture, Biological, and Environmental Statistics 3:48-61.

Rosgen, D.L. 1994. A classification of natural rivers. Catena 22:169-199.

SAS Institute. 1994. SAS User's Manual Update for Release 6.10. SAS Institute Inc., Cary, North Carolina.

Saunders, R.L., and J.H. Gee. 1964. Movement of young Atlantic salmon in a small stream. Journal of the Fisheries Research Board of Canada 21:27-36.

Schuck, H.A. 1943. Survival, population density, growth, and movement of wild brown trout in Crystal Creek. Transactions of the American Fisheries Society 73:209-230.

Shepard, B.B., J. Robison-Cox, S.C. Ireland, and R.G. White. 1996. Factors influencing retention of visible implant tags by westslope cutthroat trout inhabiting headwater streams of Montana. North American Journal of Fisheries Management 16:913-920.

Shepard, B.B., S.C. Ireland, C.C. Downs, and R.G. White. 1998a. Population dynamics and demographics of westslope cutthroat trout Oncorhynchus clarki lewisi inhabiting isolated headwater tributaries of Montana. Final Report under contract INT-RJVA-92682 (Part 2) to the USDA Forest Service, Rocky Mountain Research Station, Boise, Idaho by the Montana Cooperative Fishery Research Unit, Montana State University, Bozeman, Montana.

Shepard, B.B., M. Taper, and R.G. White. 1998b. Influence of abiotic and biotic factors on abundance of stream-resident westslope cutthroat trout Oncorhynchus clarki lewisi in Montana streams. Final Report under contract INT-RJVA-92682 to the U.S.D.A. Forest Service, Rocky Mountain Research Station, Boise, Idaho by the Montana Cooperative Fishery Research Unit, Montana State University, Bozeman, Montana.

Shetter, D.H. 1937. Migration, growth rate, and population density of brook trout in the North Branch of the Au Sable River, Michigan. Transactions of the American Fisheries Society 66:203-210.

Statistical Sciences. 1993. Splus. Seattle, Washington.

Vreeland, 1990. Random-sampling design to estimate hatchery contributions to fisheries. American Fisheries Society Symposium 7:691-707.

Zar, J.H. 1974. Biostatistical analysis. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Appendix A

Robison-Cox, J.R. 1998. A record linkage approach to imputation of missing data: analyzing tag retention in a tag-recapture experiment. Journal of Agriculture, Biological, and Environmental Statistics 3:48-61

final 9-1-97

A Record Linkage Approach to Imputation of Missing Data: Analyzing Tag Retention in a Tag–Recapture Experiment

James F. Robison-Cox

Department of Mathematical Sciences

Montana State University–Bozeman

Bozeman, MT 59715

Author footnote:

James Robison-Cox is an Assistant Professor of Statistics, Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717. The author thanks Brad Shepard for bringing up the problem and making the data available. Data collection was funded by Intermountain Research Station, Forest Service, USDA; Montana Department of Fish, Wildlife and Parks; Montana Cooperative Fisheries Research Unit, Montana State University; and the Bureau of Land Management Headwaters and Judith Resource Areas. Thanks are also extended to Robert Boik, Marty Hamilton and Fred Ramsey for helpful discussions and suggestions. The data may be obtained by emailing the author at jimrc@math.montana.edu.

Abstract

Record linkage has been used to combine records in separate files which relate to the same individual or household. In this article, modifications of the typical record linkage algorithm are made for a setting where links are "competitive", that is, a record can only be linked to once. The goal of the linkage is to recover a missing covariate for use in the logistic regression of tag retention on fish length at time of tagging. The imputation process necessarily adds measurement error to the covariate, biasing logistic regression coefficient estimates toward zero. The extent of the bias is estimated through simulations, and adjustments are made to account for the shrinkage.

The data analyzed came from a study of tag retention for visual implant tags in cutthroat trout where the measurements taken at time–of–tagging were unavailable for fish which lost the single tag.

Key words: bias adjustment; logistic regression; mark–recapture; measurement error; statistical matching.

1 INTRODUCTION

Record linkage, or statistical matching, has been used in applications such as mortality studies (Newcombe 1988) and census followup studies (Scheuren and Winkler 1993). The tag–recapture data analyzed herein is a biological application which requires some modification of record linkage algorithms in common usage for survey data. I will begin with some motivation for the analysis of this particular tag retention data.

Estimating total size of a closed population is a common problem for biologists, and can be studied using a mark–recapture protocol, however when population dynamics are of interest, marks or tags which uniquely identify individuals are required. Use of individually coded tags allows estimation of survival rates, capture probabilities, and movement. The focus of this article is on tag–recapture data reported in Shepard, Robison-Cox, Ireland, and White (1996) and described below. Analysis is complicated by the fact that the study used marks and single tags, rather than applying two tags to each fish as has been recommended for tag retention studies (Seber 1982). Double–tagged fish are less likely to lose the doubled identifying codes, so one is almost always able to relate time–of–tagging information to the tag–retention outcome in order to build a predictive model.

In order to impute the missing time–of–tagging information for fish which had lost tags, a record linkage procedure was implemented to estimate links between fish which lost tags and fish originally tagged. Record linkage involves comparing an inquiry record to a set of database records and deciding whether the inquiry matches one of the database records. For tag–recapture data, the object was to decide which tagged fish were later found without the tag. Before describing common uses of record linkage in section 3, the tag–recapture

data will be described (section 2) so that I can identify differences between the common setting and the dataset under analysis. Section 4 deals with estimation of the probability of tag retention where streams are considered to contribute random effects, and section 5 is a discussion of simulations and adjustments to the probability estimates. Discussion in section 6 is followed by an appendix which summarizes the notation used.

2 DATA DESCRIPTION

Fisheries biologists are interested in the population dynamics of westslope cutthroat trout, *Oncorhynchus clarki lewisi*, because the range and abundance of cutthroat trout have decreased dramatically (Liknes and Graham 1988, McIntyre and Rieman 1995) and this subspecies may qualify for listing under the Endangered Species Act across all or part of its range. These trout are native to western Montana, northern Idaho, and a few drainages in Washington, but have been displaced in many streams by introduced species. Data analyzed in this paper were collected in 1993 and 1994 from 9 tributary drainages of the upper Missouri River and two drainages of the upper Clark Fork River in western Montana. Beginning in summer 1993, fisheries biologists captured fish using electroshock, measured each fish and both marked (by clipping the adipose fin) and tagged (with a visual implant tag) each fish of length 120 mm or greater. For each reach of each stream, two resamplings were completed in 1993 and two or three in 1994. For each sample after the first, information was recorded on each fish captured, and in 1993 new tags were inserted in fish which had not been previously tagged, or had lost a tag.

Mark–recapture studies are employed by biologists wishing to estimate population sizes,

but biologists often desire more information on survival rates and movement patterns. To learn about population dynamics and movement, individuals must be identifiable, so tag–recapture studies are needed. The study discussed herein is part of an effort to evaluate the success, i.e., the retention probability, of a new type of tag. Double–tagging was not done in this study due to restrictions on funding and the time required to insert the second tag. (Fish are anesthetized for measurement and tagging, and the time of anesthetization must be minimized.) The study is more informative than a simple mark–recapture study, but less informative than a double–tag–recapture study. Since an objective of this study was to *predict* tag retention as a function of length and other covariates, the relevant variables are those which were measured at the time–of–tagging, and most of the statistical analysis in this article deals with imputing missing time–of–tagging information and estimating the effect of such imputation on the estimates of the coefficients relating length to tag retention.

The visual implant tags used in this study are small (2.5 mm $\times$ 1 mm $\times$ 0.1 mm) individually coded pieces of mylar which are inserted under clear skin to identify fish (Haw, Bergman, Fralick, Buckley, and Blankenship 1990). Other tags which protrude from the fish may be lost through wear or snagging on vegetation (Barrowman and Myers 1996), but loss of the visual implant tag generally occurs soon after insertion due to the tag working out of the entry site before the wound heals (Kincaid and Calkins 1992). The loss mechanism suggests the assumption that tag loss is instantaneous and not related to time–at–large. Prior experimentation has shown that visual implant tag loss is related to age of fish and to length, a commonly used surrogate for age, with tag retention being more common in larger, older fish (e.g. Niva 1995).

The data consisted of three datasets with different variables in each. Important covariates include length, date and location within stream.

TR *(tag–recaptured)*: Fish which were tagged and recaptured with tags. The response, *tag–retained*, is 1; covariates are available for both time–of–tagging and time–of–recapture. A total of 399 fish in 11 drainages were recaptured with tags, and some of these were recaptured several times.

TL *(tagged–lost)*: Fish which were tagged, but never recaptured with tags. These fish could have died, moved out of the study region, been missed in subsequent samples, or have been recaptured without a tag, so the response is missing; covariates are available only for time–of–tagging. There were 1642 fish in this dataset.

NTR *(no–tag–recaptured)*: Fish which were recaptured with clipped adipose fin and without tag, a subset of the *tagged–lost* fish. Response *tag–retained* is 0; covariates are available for time–of–recapture only. Altogether 408 fish were recaptured without tags, and since tags were reapplied at recapture, these fish are considered unique except for possible overlap in the 1994 samples.

The three datasets were further broken down by stream since streams are isolated from each other, and stream effects were expected to be strong. In seven of the eleven streams studied, at least 10 fish were in the NTR dataset, and these seven datasets were used for the complete analysis. The individual streams were chosen (not randomly sampled) as sites with large populations of cutthroat trout from a total of roughly 300 western Montana streams known to contain the subspecies, so results obtained herein can be extrapolated to the other

streams only if one is willing to make large assumptions about the representativeness of the streams chosen.

All analyses were carried out in S–plus (Statistical Sciences 1993).

3 RECORD LINKAGE

3.1 Review

A review of recent record linkage literature begins with Fellegi and Sunter (1969) who gave theoretical foundations for record linkage as a form of hypothesis testing. Applications involving linkage estimation include Newcombe (1988) who deals with a mortality study in which the relevant database is a file of death certificates, and Copas and Hilton (1990) who detail matching an exit form filled out as a foreign national leaves the United Kingdom with the same form which was filed out upon entry. Yesilcay (1993) refers to other applications in which records from two files must be matched, including references to IRS documents in which tax forms are matched. Scheuren and Winkler (1993) discuss the use of record–linked census followup datasets, and emphasize the effect of record linkage on estimation in a regression setting.

The tag–retention data differs from examples of Newcombe, Copas and Hilton, and Yesilcay in several ways:

- Inquiry records (NTR fish) **must** have a match in the TL database. Since all NTR fish had been fin–clipped, each was once tagged and a false match to a fish which had never been tagged is not possible.

- It is generally not possible for a TL fish to become two different NTR fish, so links are "competitive". Exceptions did occur for the last samples in 1994 which could include fish captured in the first 1994 sample (no new tags were applied in 1994). In the analysis, the early and late 1994 NTR samples were analyzed separately so that the competitive links structure was valid.

- As in Scheuren and Winkler, the main concern in this article is to assess the effects of imputing links on regression (in our case, logistic regression) coefficient estimates in terms of bias and variance.

The innovations presented herein are in dealing with the competitive nature of the links and in recognition that, for logistic regression, inferences depend only on the **set** of TL fish chosen, not on which individuals are linked (see the end of Section 4). Simulations provide an estimate of the biases introduced through the record linkage process.

3.2 Linkage estimation for tag retention data

Record linkage was used because the object was to recover time–of–tagging information in the TL dataset to use with NTR fish known to have lost the tag. An EM algorithm for imputing the missing TL responses was considered, but with binary response for tag retention, the "E" step of replacing a missing response with its expectation violates the binary response assumption. In record linkage algorithms, links are generally imputed using a maximum likelihood criterion whereby distributions of covariates (or data fields) are estimated, and links are chosen to maximize the product probability for the covariates assuming fields are independent (Newcombe 1988). In our case, the available covariates were

length, stream location, and dates, each measurement taken on TL fish at time–of–tagging and on **NTR** fish at time–of–recapture. Probability estimation was carried out by using the 399 TR fish to estimate the distributions for movement and growth, and, assuming that movement and growth to be mutually independent and independent of tag–retention, then applying these estimated distributions to the TL and NTR datasets. The three covariate distributions are discussed individually, and then they are combined into a joint distribution for a set of links.

Growth. Biologists have observed that fish growth occurs when water temperatures are above a certain minimum, and that growth rates vary with temperature and availability of food. A model in common use for growth over time assumes constant growth rate over some prespecified (warm) season. The model used in this analysis was based on a linear regression of square root growth on time–at–large and was generalized to allow simultaneous estimation of the length of the non-growth season (winter) and the growth rate. For streams, $s = 1, \ldots, 11$ regressions were performed using the following model for mean square root growth:

$$\mu_{ijs} = \beta_{1s}(t_j - t_i) - \beta_{2s} I_W(t_i, t_j) \tag{1}$$

Where $I_W(t_i, t_j) = 1$ if the interval between the two dates included January 1, 1994 and

Table 1: Estimated Growth Rates and Winter Lengths

Stream	Collar	Cottnwd	Douglas	Geyser	Gold	Halfway	Jerry	Mcvey	Soap
Growth	0.027	0.025	0.038	0.033	0.030	0.021	0.020	0.073	0.041
Std err	0.0016	0.0032	0.0054	0.0025	0.0058	0.0032	0.0036	0.013	0.0052
Winter	159.2	166.3	221.9	230.7	218.6	179.3	170.1	243.9	181.6

was 0 otherwise. The first coefficient, β_{1s}, is the daily increase in square root length, and the

second, β_{2s} represents the unattained growth due to winter. Data points and the estimated growth lines are shown in Figure 1. Use of square roots made the assumptions of normal errors and constant variance tenable, so the MLE for length of winter on stream s is $\hat{\beta}_{2s}/\hat{\beta}_{1s}$. These estimates are shown in Table 1.

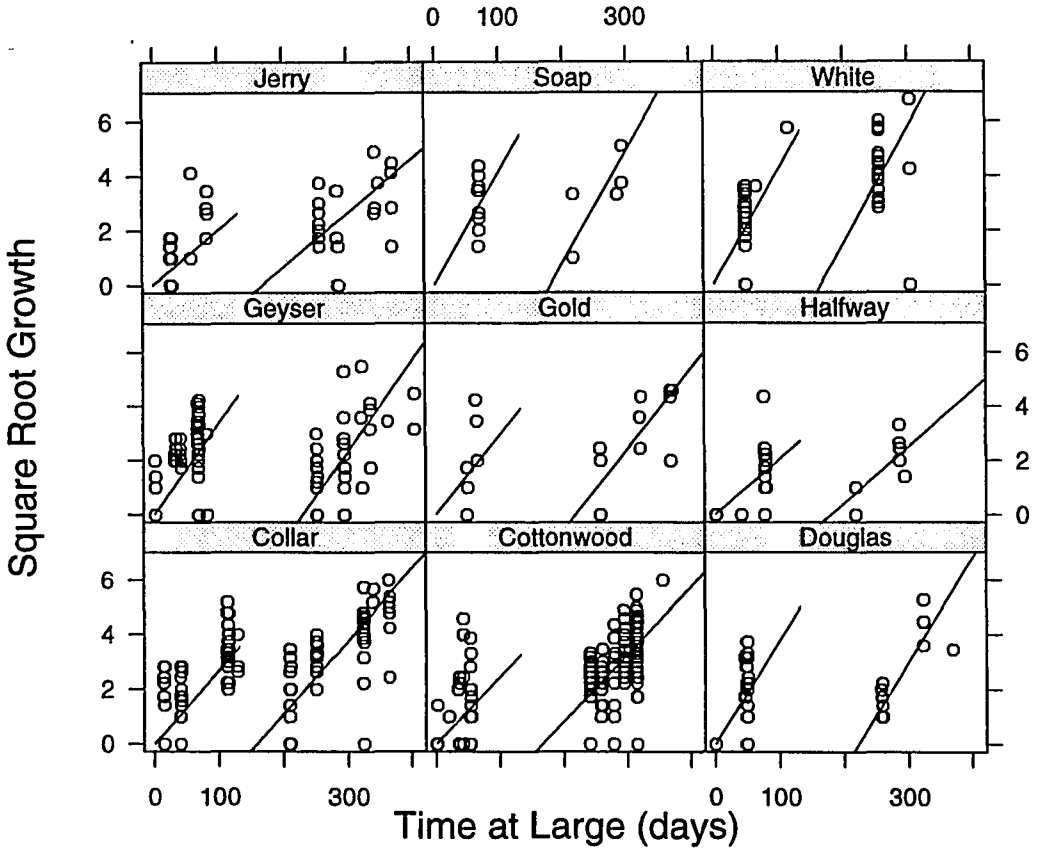


Figure 1: Square root growth ($\sqrt{\text{mm}}$) as a function of time–at–large. Times less than 150 days represent fish released and recaptured within 1993, larger times indicate fish released in 1993 and recaptured in 1994. Of the 399 recaptured fish, the 109 which were recaptured more than once are represented as multiple points.

The predictive distribution for square root growth, $\sqrt{g_{ijs}}$, on stream s over days t_i to t_j based on slope estimates in Table 1 was taken as normal with mean and variance

$$\hat{\mu}_{ijs} = \hat{\beta}_{1s}(t_j - t_i) - \hat{\beta}_{2s} I_W(t_i, t_j)$$

$$\hat{\sigma}^2_{ijs} = \hat{\sigma}^2_g + \hat{\text{var}}(\hat{\beta}_{1s})(t_j - t_i)^2 + [\hat{\text{var}}(\hat{\beta}_{2s}) - 2\hat{\text{cov}}(\hat{\beta}_{1s}, \hat{\beta}_{2s})(t_j - t_i)]I_W(t_i, t_j).$$

The extra component of variance due to estimation of a new growth, $\hat{\sigma}^2_g$, was estimated over all streams as 1.39. To compute the probability that tagged–lost fish i would grow g_{ij} mm to the length of no–tag–recaptured fish j in time $t_j - t_i$, the continuity corrected (length was measured to the nearest mm) normal probability was used:

$$P_{Gs}(g_{ij}) = \Phi\left(\frac{\sqrt{g_{ij} + .5} - \hat{\mu}_{ijs}}{\hat{\sigma}_{ijs}}\right) - \Phi\left(\frac{\sqrt{g_{ij} - .5} - \hat{\mu}_{ijs}}{\hat{\sigma}_{ijs}}\right). \quad (2)$$

Movement. The probability distribution for distance moved was estimated from the frequency of observed moves in the tag–recaptured fish. Roughly 70% of the fish did not move, but two fish moved more than 9 km. A smoothed empirical pmf was obtained by assuming the moves were symmetric about zero (upstream or downstream are equally likely) and that moves were not dependent on time–at–large ($r = 0.075$) or length ($r = 0.08$). Table 2 lists estimated probabilities of non–negative moves

Table 2: Distance moved (km) probabilities estimated and smoothed over all streams. Negative moves have identical probabilities.

m_{ij}	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	≥ 1
$P_M(m_{ij})$	0.69	0.063	0.036	0.018	0.012	0.010	0.007	0.003	0.003	0.002	0.001

Dates. Since a fish could not have been recaptured before it was originally tagged, probabilities of i matching j were taken to be zero if tag date of fish i was later than recapture date of fish j. Indicator function I_D was used to ensure that the probability of tagged–lost fish i matching no–tagged–recaptured fish j was zero if $t_j \leq t_i$.

Combining covariate information. Record linkage methods commonly assume that

fields are independent for matching purposes. With this data, growth and movement were assumed to be independent, an assumption which was not contradicted by the data. (Scatterplots revealed little structure, $r = 0.08$ and non-parametric tests for zero correlation gave p-values over 0.10.) The independence assumption allows computation of probability P_{ijs} of tagged–lost fish i matching no–tagged–recaptured fish j in stream s based on growth, movement, and difference in dates as:

$$P_{ijs} = P_{Gs}(g_{ij}) \times P_M(m_{ij}) \times I_D(t_j > t_i) \tag{3}$$

Note that the effect of move is very strong in the above product, since moves of zero are roughly 10 times more probable than moves of $\pm.1$ km. One can adjust the relative importance of movement versus growth by varying the scale of the normal distribution used in $P_{Gs}(g_{ij})$, but changes seemed to have little effect on the matching in simulations described below.

Linkage Estimation For simplicity, the stream subscript will be dropped at this point; what follows assumes the stream is specified. A matrix of match probabilities was created with $j = 1, \ldots, k$ rows and $i = 1, \ldots l$ columns by making each element of the matrix P_{ij}, the probability of link L^i_j. If only individual links were needed, the maximum likelihood individual link for no–tag–recaptured fish i' would be the tagged–lost fish represented by j^*, the column index of the largest entry in row i' of the match probability matrix. However, the object was to find the **set** of TL fish corresponding to the set of **NTR** fish, so a **maximal set** of links is needed. Any set of links is a sample taken without replacement from the set of

possible links, and the order of the sample is an essential part of the probability calculation. Assuming competitive links, so that L^i_j and $L^i_{j'}$ are not both possible unless $j = j'$, the conditional probability of linking i_2 to j_2, given i_1 links to j_1 (event $L^{i_2}_{j_2}|L^{i_1}_{j_1}$) is $P_{i_2j_2}/(1-P_{i_2j_1})$ which extends to give the probability of the ordered set of links $\{L^{i_1}_{j_1}, L^{i_2}_{j_2}, \ldots L^{i_k}_{j_k}\}$ as

$$P_{i_1j_1} \times \frac{P_{i_2j_2}}{(1-P_{i_2j_1})} \times \cdots \times \frac{P_{i_kj_k}}{(1-\sum_{r=1}^{k-1} P_{i_kj_r})}. \tag{4}$$

To find the unordered probability of a set of links, we would have to sum these probabilities over all $k!$ possible orderings, which is not feasible. Instead, to compare two different sets of links, random permutations of link orders were generated, for each ordering the probability in (4) was computed, and sums or means of logs of the probabilities in (4) were compared. In the data examined here, the effect of ordering was small. Samples of 1000 random orderings from sets of links gave distributions of log likelihoods which were symmetric (mean and median coincided and first and third quartiles were evenly spaced about the median), with small to moderate spread (means ranged from -37 to -221 over the seven streams with standard deviations ranging from .51 to 1.97).

Two methods were used to estimate the "best" set of links. First, a method I will refer to as the "Mratio" approach, found a set of ordered links by maximizing each choice in order. Under this scheme, for each row of the match probability matrix, a ratio was calculated of the largest match probability divided by the next largest match probability. Then rows were compared to find the highest ratio. If row i' had the highest ratio, the NTR fish i' was linked to its best j' of the TL fish, row i' and column j' were removed, and the process was repeated for the remaining rows. Rather than looking at all NTR fish for the link

with highest match probability, the procedure found the link which contributed most to the ordered probability in (4). As would be expected with a stepwise maximization method, the order of the set of links was generally not optimal, for by randomly reordering the links 500 times, a better ordering was found for Mratio links from six of the seven streams.

Secondly, random orderings $(i_{p_1}, \ldots, i_{p_l})$ of the rows of the probability matrix were generated and for each ordering, at stage $q = 1, \ldots, l$, the best match, j' in row $i' = i_{p_q}$ was selected as a link between NTR fish i' and TL fish j'. This method of random choice generated the links in a random order without regard to the probability in (4), so was expected to work only by brute force, with large numbers of random choices. To compare the two methods, 1000 random link choices were made and the ten sets with largest ordered probabilities were compared to the set of Mratio links. In six of the seven streams, the links picked by the the Mratio method had higher mean log ordered probability over 500 random permutations of orderings than all ten of the "best" randomly chosen links. In the seventh stream, Halfway Creek, the Mratio links had mean log ordered probability of -48.85 while the best random links had means from -48.52 to -48.20; each a slight improvement over the Mratio links. For simulations, the Mratio method was used because it generally gave higher mean log probabilities over random orderings, and because it was much faster.

An important observation about the linkage for this data is that for estimating effects of covariates on tag retention through logistic regression, we can not only disregard the order of links, but also the individual link estimates (which TL fish links to which NTR fish). The set of TL fish which were linked is the critical imputation, since it is those fish which were given response of tag–retention = 0 and included in the logistic regression dataset. For

Halfway Creek, the Mratio links were outdone by the random choices of links, but the sets of TL fish considered were very similar. The ten highest probability sets of random links (32 links) each differed from the Mratio links in only one or two fish choices out of 158 TL fish possible. The importance of only the set of TL fish chosen for links also allowed combining of two sets of links, those from the data collected in 1993 and early 1994 and that from late 1994 where some links could be shared between the early and later 1994 samples. Each set of links was estimated under the assumption of competitive links, then the two sets were merged to give the set of TL fish imputed to have lost tags.

In order to check the accuracy of the linkage estimation, another regression of square root growth on time–at–large was performed, this time combining the known (tag-recaptured) data with the imputed data from Mratio links. Examination of residuals showed that record linkage may have over–fit the data, since standard deviations were 0.83 and 1.15 for residuals of imputed and known data points respectively. Coefficient estimates showed little change, at most changing by 1 standard error of the slope estimate, usually by only 0.2 standard errors (the mean of shift/(original standard error) over the 7 streams). When merging the set of links for late 1994 recaptures with earlier recaptures, growth and time–at–large were taken only from the later recaptures since these fish could have been recaptured in early 1994 as well, but the later dates allowed the fish to grow more, and hence contained more information about growth rates.

4 ESTIMATION OF PROBABILITY OF TAG RETENTION

Binomial response data has been studied extensively, and adjustments to logistic regression for repeated measures over time have been addressed in several ways (Schall 1991, Stiratelli, Liard, and Ware 1984, and Zeger, Liang, and Albert 1988). Data treated in the references listed are all cases where a few repeated measurements were recorded on each of many individuals, in contrast to the tag–retention data analyzed here where the number of streams (individuals) is small and the number of fish (repeated measurements) within each stream is fairly large, allowing individual logistic regressions on each stream. The analysis would be most useful if it provided a marginal, or population average, estimate of the probability of tag retention as a function of length which would hold for all fish, regardless of the stream. In terms used by Zeger, et al. (1988), the desired estimates are for a population–average model, rather than subject–specific models, so some means of averaging together the results of the individual stream models is needed. The averaging is commonly done by assuming a normal distribution for random effects, but with only seven streams, one has little basis for assuming that the normal distribution (or any other candidate distribution) is appropriate, so I have used the empirical distribution, $\hat{F}$ which gives equal weight to each stream. The population–averaged probability of retention, as a function o length, is then simply the average of the stream–specific probabilities at each given length.

A glance at Figure 2 shows that in one stream the relationship between length and probability of tag retention is reversed. The data from Halfway Creek is anomalous in that only two fish were recaptured at lengths greater than 180 mm, and both had lost tags. A waterfall in Halfway Creek allows fish to move downstream, exiting the area studied, but

prevents them from coming back in, and apparently few large fish stay in the study area. Since this anomaly is explainable in terms of the stream topography, Halfway was left out of the averaging process. After adjustments discussed in the next section, the final model is based on averaging the mean values for each of the other six streams at each length.

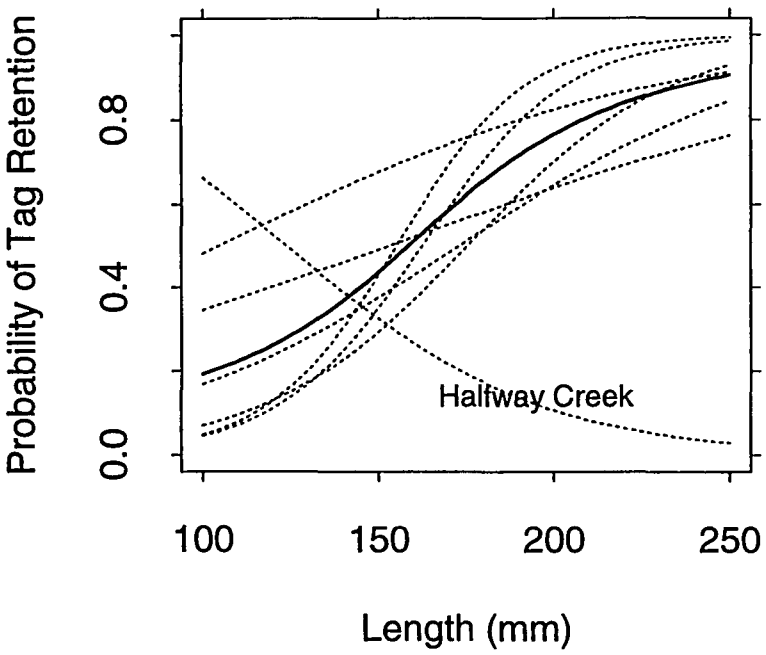


Figure 2: Stream specific (dotted lines) estimates of tag retention probability as a function of length adjusted for bias due to measurement error. Halfway Creek, the one unusual stream is labeled. Solid line is the population–averaged estimate omitting Halfway Creek.

5 SIMULATIONS AND ADJUSTMENTS

Simulations were carried out to examine properties of the link matching process and to address questions about bias and variance inflation of regression coefficients due to use of estimated time–of–tagging lengths instead of the true, unknown lengths. For each of the

seven streams in Table 3 a new dataset was created using all n of the TR fish and all l of the TL fish, completing the missing data for the TL by generating:

a recapture date,	a random draw from sampling dates later than the tagging date,
a recapture location,	tag location plus a draw from (-1,1)km using the move probabilities from Table 2,
a recapture length,	a random drawn from the predictive normal distribution estimated in section 3 using known tagging date and recapture date generated above.

Table 3: Logistic regression coefficients for tag retention on length from original GLM and simulations.

Stream	Original Estimate	Simulation Mean	% Shrink (Sim/Orig)	Adjusted Estimate	Original Std. Err	Simulation Std. Dev
Collar Gulch	0.0103	0.0037	35.7	0.0161	0.0084	0.0076
Cottonwood Creek	0.0350	0.0163	46.5	0.0544	0.0063	0.0052
Douglas Creek	0.0314	0.0190	60.3	0.0488	0.0136	0.0126
Geyser Creek	0.0141	0.0096	68.4	0.0219	0.0047	0.0045
Halfway Creek	-0.0179	-0.0209	116.0	-0.0280	0.0153	0.0203
Jerry Creek	0.0077	0.0050	64.3	0.0120	0.0107	0.0106
Soap Creek	0.0221	0.0144	65.2	0.0344	0.0140	0.0131

The partially–invented data was used as the original population of fish. For each simulation run on a given stream, an artificial population of fish was created by first picking n fish at random to be the TR fish based on probabilities of retention as estimated from the coefficients $\hat{\gamma}_0$ and $\hat{\gamma}_1$ in Table 3, column 1. Using Y for the tag retention response and Z for length,

$$\text{logit}(P(Y = 1|Z = z)) = \hat{\gamma}_0 + \hat{\gamma}_1 z.$$

Of the remaining l fish, k were selected as the NTR fish using empirical marginal probabilities estimated across all streams, where probability of recapture is modeled as a function of length z.

$$\text{logit}(P(\text{NTR}|Z = z)) = 1.314 - 0.0066z$$

The simulation of data was repeated 100 times and for each simulated dataset the Mratio linkage algorithm described in Section 3 was used to impute missing data. A logistic regression was fit to the data from each simulation and the coefficient estimates were saved for comparison with the starting values in column one of Table 3.

Table 3 shows the important results of the simulations by comparing means and standard deviations of estimates from the simulations with the original estimates. The last two columns are quite similar and show that the variance estimates of logistic slopes from the generalized linear model are quite accurate despite the measurement error introduced by record linkage. Record linkage has introduced substantial bias, as demonstrated by comparing the glm estimate in column one with the mean slope from 100 simulations in column two. The third column is the percent shrinkage of the simulation mean as compared to the original (column 1) value. The coefficients from simulation are biased toward 0, the means typically being 65% of the original values with Halfway Cree again being anomalous. Such bias is expected since the record linkage estimation has added measurement error to the tagging–lengths. Neter, Maynes, and Ramanthan (1965) recognized that mismatching causes bias of regression coefficients toward zero, and Scheuren and Winkler (1993) generalized those results. I choose to adjust the original estimates by dividing by 0.643, the median of the shrinkages. Intercept terms were modified to avoid change in the estimated EL50,

length for which 50% of the fish would retain their tags. Figure 2 illustrates the variation in adjusted coefficients, plotting the predictive curve for each stream as a dashed line, and the average of the six curves (excluding Halfway) as a solid line.

The mean of the logistic curves is not itself a logistic, for the curvature is not quite symmetric in the left and right tails. The best fitting linear logistic curve is $\text{logit}(\mu_z) = -4.13+0.0261z$, which is off by at most .017 in the probability scale. A closer approximation to the population average curve is given by the cubic logistic, $\text{logit}(\mu_z) = -0.747-0.0397z+$ 4.08e-05z^2 – 8.09e-07z^3, which is never more than .005 from the average curve for lengths from 100 to 250 mm.

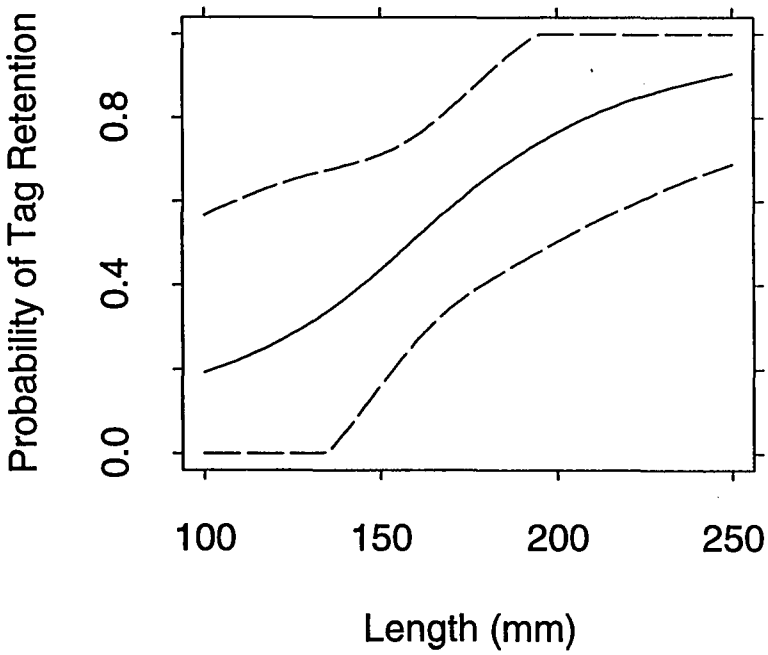


Figure 3: Final population–averaged estimate of tag retention probability with 2 standard error confidence band.

Variances of the population averaged mean for different lengths are the sum of a variance

component for between–stream variation and the sum of estimation variances from the six different fitted values divided by 36. As one might guess, the between–stream component of variation is much larger than the component from estimation of stream parameters; it accounts for roughly 90% of the variance. In Figure 3, the population average is again plotted along with 2 standard error bands, approximate 95% confidence bounds for the estimation of the mean probability of tag retention given a particular length.

As a final point about the simulations, it is also of interest to examine the effectiveness of a record linkage strategy. In this case, the important criterion is not the number of correct individual links, but, rather the percent of pseudo–NTR fish which were properly identified. The set–wise linkage success rates ranging from 38 to 61% (means over 100 simulations) were an improvement over random drawing of k NTR's from the l TL fish where expected success rates k/l ranged from 17 to 34% over the seven streams.

6 DISCUSSION

The single tag with mark protocol greatly complicated analysis as compared to double tagging studies, but record linkage did provide a reasonable method of imputation for these data and will be useful in other settings where links are competitive. Estimation of links cannot be done without errors, so measurement error becomes an source of substantial bias in the regression coefficients just as Scheuren and Winkler (1993) discussed in a standard regression setting. Simulations also showed that the standard errors of coefficient estimates reported by glm were reasonable. Neter et al. (1965) were able to estimate bias of slope as a function of the probability of correct match (assumed constant over all individuals),

and perhaps their results are extensible to the case of non-constant match probabilities. However, the lack of a visible relationship between the match success rates and the bias rates in Table 3 makes me doubt that a theoretical relationship between the two exists.

The next step in handling tag–recapture data with imputed links is to use the data to estimate population abundance, capture probabilities, and survival probabilities. Since an assumption of closed population is not tenable, estimating abundance is not straightforward, but may be possible by adjusting counts to reflect tag loss. Capture and survival probability estimates may suffer too much from variation of link imputation to give useful results. These issues are being investigated.

Appendix

Notation, approximately in order of appearance in the text:

$s = 1, \ldots, 11$	stream number.
n_s	number of tag–recaptured (TR) fish.
k_s	number of no–tag–recaptured (NTR) fish.
$j = 1, \ldots, k_s$	an index for NTR fish.
l_s	number of tagged–lost (TL) fish ($l \geq k$).
$i = 1, \ldots, l_s$	an index for TL fish.
g_{ijs}	growth needed for TL fish i to become NTR fish j in stream s.
$I_W(t_1, t_2)$	indicator function for winter, 1 if January 1, 1994 is between dates t_1 and t_2.
β_{1s}	daily increase in square root length.
β_{2s}	unattained growth due to winter.
L_j^i	event that TL fish i is recaptured as (or linked to) NTR fish j.
m_{ij}	movement (in km) needed for TL fish i to become NTR fish j.
$P_{Gs}(g, \delta_t)$	probability a fish grows g mm in δ_t days.
$P_M(m)$	probability a fish moves m km.
$t_j - t_i$	number of days from tagging TL fish i to recapture of fish j.
$I_D(t_j > t_i)$	indicator function for tag date preceding recapture date.
γ_{0s}	logistic intercept for tag retention probability on stream s.
γ_{1s}	logistic slope (on length) for stream s.
z_i	length of fish i.

References

Barrowman, N. J. and Myers, R. A. (1996), "Estimating Tag–shedding Rates for Experiments with Multiple Tag Types", *Biometrics*, 52, 1410–1416.

Copas, J. P., and Hilton, F. J. (1990), "Record Linkage: Statistical Models for Matching Computer Records", *Journal of the Royal Statistical Society*, Ser. A, 153, 287–320.

Fellegi, I. P. and Sunter, A. (1969), "A Theory of Record Linkage", *Journal of the American Statistical Association*, 64, 1183–1210.

Haw, F., Bergman, P. K., Fralick, R. D., Buckley, R. M., and Blankenship, H. L. (1990), "Visible Implanted Fish Tag", *American Fisheries Society Symposium 7*, 311–315.

Kincaid, H. L. and Calkins, G. T. (1992), "Retention of Visible Implant Tags in Lake Trout and Atlantic Salmon", *Progressive Fish–Culturist*, 54, 163–170.

Liknes, G. A. and Graham, P. J., (1988), "Western Cutthroat Trout in Montana: Life History, Status and Management", *American Fisheries Society*, 4, 53–60.

McIntyre, O. P. and Rieman, B. E. (1995), "Westslope Cutthroat Trout" in *Conservation Assessment for Inland Cutthroat Trout*, ed. M. K. Young, General Technical Report RM-256, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, 1–15.

Neter, J., Maynes, E. S., and Ramanthan, R. (1965), "The Effect of Mismatching on Measurement of Response Errors", *Journal of the American Statistical Association*, 60, 1005–1027.

Newcombe, H. B. (1988), *Handbook of Record Linkage* Oxford: Oxford University Press.

Niva, T. (1995), "Retention of Visible Implant Tags by Juvenile Brown Trout, *Salmo Trutta*", *Journal of Fisheries Biology*, 46, 997–1002.

Schall, R. (1991), "Estimation in Generalized Linear Models with Random Effects", *Biometrika*, 78, 719–727.

Scheuren, F. and Winkler, W. E. (1993), "Regression Analysis of Data Files that are Computer Matched", *Survey Methodology*, 19, 39–58.

Seber, G. A. F. (1982), *The Estimation of Animal Abundance*, New York: Macmillan. correction: (1981) *Biometrika*, p 211.

Shepard, B. B., Robison-Cox, J., Ireland, S. C., and White, R. G. (1996), "Factors Influencing Retention of Visible Implant Tags in Westslope Cutthroat Trout Inhabiting Headwaters Streams of Montana", *North American Journal of Fisheries Management*, 16, 913–920.

Statistical Sciences (1993), *S-PLUS User's Manual*, version 3.2, Seattle: StatSci, a division of MathSoft, Inc.

Stiratelli, R., Liard, N., and Ware, J. H. (1984), "Random–effects Models for Serial Observations with Binary Response", *Biometrics*, 40, 961–971.

Yesilcay, Y. (1993), "The BOYS Algorithm for Determining Optimum Matching Rules", *Computational Statistics and Data Analysis*, 16, 219–234.

Zeger, S. L., Liang, K. Y., and Albert, P. S. (1988), "Models for Longitudinal Data: A Generalized Estimating Equations Approach", *Biometrics*, 44, 1049–1060.

Appendix B

Data Structure for Tag Data

TAGKM.DBF contains tag and recapture data for all fish tagged, whether recaptured or not.

Structure for table D:\TORMS\TAGKM.DBF
Table type DBASE
Number of records 3955
Last update 10/19/98

Field	Field Name	Type	Length	Dec	Index
1	LOCATION	CHARACTER	25		N
2	SEC_NUMBER	NUMERIC	4	1	N
3	SPEC_NUM	CHARACTER	3		N
4	DATE	DATE	8		N
5	LENGTH_MM	NUMERIC	6	1	N
6	WEIGHT_GM	NUMERIC	4		N
7	TAGTYPE	CHARACTER	2		N
8	TAGCOLOR	CHARACTER	2		N
9	TAGNUMBER	CHARACTER	6		N
10	TAGRATING	CHARACTER	1		N
11	FINCLIPPED	CHARACTER	2		N
12	STREAMKM	NUMERIC	7	2	N
13	RECAP1DATE	DATE	8		N
14	RECAP1LENG	NUMERIC	6	1	N
15	RECAP1WEIG	NUMERIC	4		N
16	RECAP1KM	NUMERIC	7	2	N
17	RECAP1TRIB	LOGICAL	1		N
18	RECAP2DATE	DATE	8		N
19	RECAP2LENG	NUMERIC	6	1	N
20	RECAP2WEIG	NUMERIC	4		N
21	RECAP2KM	NUMERIC	7	2	N
22	RECAP2TRIB	LOGICAL	1		N
23	RECAP3DATE	DATE	8		N
24	RECAP3LENG	NUMERIC	6	1	N
25	RECAP3WEIG	NUMERIC	4		N
26	RECAP3KM	NUMERIC	7	2	N
27	RECAP3TRIB	LOGICAL	1		N
28	RECAP4DATE	DATE	8		N
29	RECAP4LENG	NUMERIC	6	1	N
30	RECAP4WEIG	NUMERIC	4		N
31	RECAP4KM	NUMERIC	7	2	N
32	RECAP4TRIB	LOGICAL	1		N
33	RECAP5DATE	DATE	8		N
34	RECAP5LENG	NUMERIC	5	1	N
35	RECAP5WEIG	NUMERIC	5	1	N
36	RECAP5KM	NUMERIC	7	2	N
37	RECAP5TRIB	LOGICAL	1		N
38	RECAP6DATE	DATE	8		N
39	RECAP6LENG	NUMERIC	5	1	N
40	RECAP6WEIG	NUMERIC	5	1	N
41	RECAP6KM	NUMERIC	7	2	N
42	RECAP6TRIB	LOGICAL	1		N
** Total **			227		

Field name	Explanation
LOCATION	Stream name
SEC_NUMBER	Number of sample section
SPEC_NUM	Species code: 001=rainbow trout; 003=brook trout; 004=brown trout; 011=hybrid between westslope cutthroat and rainbow trout; 012= westslope cutthroat trout; 058=largescale sucker; 085=mountain whitefish
DATE	Date fish was tagged
LENGTH_MM	Fork length measured in mm
WEIGHT_GM	Weight in grams
TAGTYPE	Tag type used to tag fish: FL=floy; VI=visible implant
TAGCOLOR	Tag color: BK=black; WH=white; and Y=yellow
TAGNUMBER	Alpha-numeric code on tag
TAGRATING	Tag insertion quality rating: 1=good insetion, tag could be easily read; 2=fair insertion, some tissue was torn during insertion or tag was inserted too deeply; 3=poor insertion, tissue was badly torn or tag was inserted so deeply it could not be read.
FINCLIPPED	Fin clipped as second mark (AD = adipose)
STREAMKM	Stream kilometer above mouth (or other major landmark that locates sample section along stream
RECAP#DATE	Recapture date with # replaced by number 1 through 6 to indicate recapture number.
RECAP#LENG	Length (fork length in mm) at recapture number #.
RECAP#WEIG	Weight (gm) at recapture number #.
RECAP#KM	Stream kilometer where fish was recaptured at recapture number #.
RECAP#TRIB	Logical field (Y or N, T or F) that indicates if fish was recaptured in same stream as tagged (N or F) or in a different tributary (Y or T) at recapture number #. If it was recaptured in a different tributary, the distance moved information must be obtained by reviewing the raw data.